GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 21:30:10 ; Search time 4788 Seconds
(without alignments)
10606.599 Million cell updates/sec

Title: US-10-017-621-3
Perfect score: 1745
Sequence: 1 tggaagcagcgtaaaggatg.....gttcacctgcccacttgtcc 1745

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2054640 seqs, 14551402878 residues

Total number of hits satisfying chosen parameters: 841850

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : GenEmbl:*
1: gb_ba:*
2: gb_htg:*
3: gb_in:*
4: gb_om:*
5: gb_ov:*
6: gb_pat:*
7: gb_ph:*
8: gb_pl:*
9: gb_pr:*
10: gb_ro:*
11: gb_sts:*
12: gb_sy:*
13: gb_un:*
14: gb_vi:*
15: em_ba:*
16: em_fun:*
17: em_hum:*
18: em_in:*
19: em_mu:*
20: em_om:*
21: em_or:*
22: em_ov:*
23: em_pat:*
24: em_ph:*
25: em_pl:*
26: em_ro:*
27: em_sts:*
28: em_un:*
29: em_vi:*
30: em_htg_hum:*
31: em_htg_inv:*
32: em_htg_other:*
33: em_htg_mus:*
34: em_htg_pln:*
35: em_htg_rod:*
36: em_htg_mam:*
37: em_htg_vrt:*
38: em_sy:*
39: em_htgo_hum:*
40: em_htgo_mus:*
41: em_htgo_other:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

Result No.	Score	% Query Match	Length	DB	ID	Description
1	23.8	1.4	49	9	HSHCAK	X76171 H.sapiens m
c 2	22.8	1.3	50	6	A93721	A93721 Sequence 8
3	22.4	1.3	50	6	A93722	A93722 Sequence 9
4	22.2	1.3	47	6	I84671	I84671 Sequence 5
5	22	1.3	50	6	AX159452	AX159452 Sequence
6	21.6	1.2	31	6	AX248673	AX248673 Sequence
c 7	21.4	1.2	42	6	AX182243	AX182243 Sequence
c 8	21.4	1.2	42	6	AX382049	AX382049 Sequence
9	21.4	1.2	46	6	AR032544	AR032544 Sequence
10	21.4	1.2	46	6	AR209208	AR209208 Sequence
11	21.4	1.2	46	6	I29284	I29284 Sequence 15
12	21.4	1.2	46	6	I90958	I90958 Sequence 15
13	21	1.2	31	6	AX248015	AX248015 Sequence
c 14	21	1.2	46	6	AX186238	AX186238 Sequence
c 15	20.8	1.2	46	6	A98791	A98791 Sequence 24
16	20.6	1.2	21	6	AX153998	AX153998 Sequence
c 17	20.6	1.2	45	6	AR022074	AR022074 Sequence
18	20.6	1.2	45	6	I55009	I55009 Sequence 33
c 19	20.6	1.2	45	6	I92864	I92864 Sequence 38
20	20.4	1.2	48	6	AR079723	AR079723 Sequence
21	20.4	1.2	48	6	AR081253	AR081253 Sequence
22	20.4	1.2	48	6	AR170613	AR170613 Sequence
23	20.2	1.2	40	6	AR200128	AR200128 Sequence
24	20.2	1.2	40	6	I68030	I68030 Sequence 13
c 25	20.2	1.2	45	6	AX225269	AX225269 Sequence
26	20.2	1.2	49	6	AR083818	AR083818 Sequence
27	20.2	1.2	49	9	S82032	S82032 WT1=Wilms'
c 28	20.2	1.2	50	6	AX233404	AX233404 Sequence
29	20	1.1	36	6	A07324	A07324 Synthetic D
30	20	1.1	36	6	I12501	I12501 Sequence 18
31	20	1.1	41	6	BD007098	BD007098 Targeted
c 32	20	1.1	44	6	A07325	A07325 Synthetic D
c 33	20	1.1	44	6	I12502	I12502 Sequence 19
34	20	1.1	46	6	E52011	E52011 IL-6 recept
35	19.8	1.1	46	6	AX036348	AX036348 Sequence
36	19.8	1.1	46	6	AX036350	AX036350 Sequence
c 37	19.8	1.1	48	6	BD012118	BD012118 Vitamin D
c 38	19.8	1.1	48	23	BD004595	BD004595 Vitamin D
39	19.6	1.1	39	6	AX452342	AX452342 Sequence
c 40	19.6	1.1	42	6	AR153233	AR153233 Sequence
c 41	19.6	1.1	45	6	I17261	I17261 Sequence 27
42	19.6	1.1	45	6	I47720	I47720 Sequence 19
43	19.6	1.1	50	6	AR032970	AR032970 Sequence
44	19.6	1.1	50	6	AR209634	AR209634 Sequence
45	19.6	1.1	50	6	AX199648	AX199648 Sequence

ALIGNMENTS

RESULT 1
HSHCAK
LOCUS HSHCAK 49 bp mRNA linear PRI 08-AUG-1995
DEFINITION H.sapiens mRNA for Cdk activating kinase.
ACCESSION X76171
VERSION X76171.1 GI:429096
KEYWORDS activating kinase; protein kinase.
SOURCE Homo sapiens.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 49)
AUTHORS Hall,F.L.
TITLE Direct Submission
JOURNAL Submitted (08-NOV-1993) F.L. Hall, Childrens Hospital Los Angeles,

```
            Division of Orthopaedic Surgery, 4650 Sunset Boulevard, Los
            Angeles, CA, 90027, USA
  REFERENCE   2  (bases 1 to 49)
    AUTHORS   Wu,L. and Hall,F.L.
    JOURNAL   Unpublished
  REFERENCE   3  (bases 1 to 49)
    AUTHORS   Wu,L., Yee,A., Liu,L., Carbonaro-Hall,D., Venkatesan,N., Tolo,V.T.
            and Hall,F.L.
    TITLE     Molecular cloning of the human CAK1 gene encoding a
            cyclin-dependent kinase-activating kinase
    JOURNAL   Oncogene 9 (7), 2089-2096 (1994)
    MEDLINE   94268852
     PUBMED   8208556
  FEATURES             Location/Qualifiers
       source          1. .49
                       /organism="Homo sapiens"
                       /db_xref="taxon:9606"
                       /clone="IARC EW-1"
                       /cell_line="EW-1 Ewings sarcoma cells"
                       /cell_type="Ewings sarcoma"
                       /tissue_type="tumour"
                       /clone_lib="IARC"
       gene            1. .49
                       /gene="HCAK"
       CDS             <1. .>49
                       /gene="HCAK"
                       /note="catalytic subunit; subdomain VII"
                       /codon_start=2
                       /product="protein kinase"
                       /protein_id="CAA53766.1"
                       /db_xref="GI:939888"
                       /db_xref="SPTREMBL:Q14495"
                       /translation="KWSSKTGRFWPGQIFW"
  BASE COUNT       14 a      8 c     12 g     15 t
  ORIGIN

    Query Match             1.4%;  Score 23.8;  DB 9;  Length 49;
    Best Local Similarity   72.1%;  Pred. No. 7.8e+05;
    Matches   31;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

  Qy   1010 AGAGGGGAGAGCTCAAGCTGGCTGACTTTGGCCTGGCCCGAGC 1052
            | |  ||||  || || ||||| || ||||||||||||  | |
  Db      1 AAAATGGAGTTCTAAAACTGGCAGATTTTGGCCTGGCCAAATC 43


  RESULT 2
  A93721/c
  LOCUS       A93721                    50 bp    DNA     linear   PAT 22-JAN-2000
  DEFINITION  Sequence 8 from Patent WO9732891.
  ACCESSION   A93721
  VERSION     A93721.1  GI:6741893
  KEYWORDS    .
  SOURCE      unidentified.
    ORGANISM  unidentified
            unclassified.
  REFERENCE   1  (bases 1 to 50)
    AUTHORS   Rose-John,S.
    TITLE     CONJUGATE FOR MODIFYING INTERACTIONS BETWEEN PROTEINS
    JOURNAL   Patent: WO 9732891-A 8 12-SEP-1997;
            ANGEWANDTE GENTECHNOLOGIE SYST (DE); ROSE JOHN STEFAN (DE)
  FEATURES             Location/Qualifiers
       source          1. .50
                       /organism="unidentified"
                       /db_xref="taxon:32644"
  BASE COUNT        7 a      4 c     26 g     13 t
  ORIGIN

    Query Match             1.3%;  Score 22.8;  DB 6;  Length 50;
    Best Local Similarity   66.0%;  Pred. No. 1.3e+06;
    Matches   33;  Conservative    0;  Mismatches   17;  Indels    0;  Gaps    0;

  Qy    961 CAGAAGGTGCTACACCGAGACCTCAAGCCCCAGAACCTGCTCATCAACGA 1010
            |||||  | |  | ||   ||||| | | |||||||||| | | ||  |||

  Db     50 CAGAACCTCCACCTCCAGAACCTCCACCTCCAGAACCTCCACCTCCTCGA 1


  RESULT 3
  A93722
  LOCUS       A93722                    50 bp    DNA     linear   PAT 22-JAN-2000
  DEFINITION  Sequence 9 from Patent WO9732891.
  ACCESSION   A93722
  VERSION     A93722.1  GI:6741894
  KEYWORDS    .
  SOURCE      unidentified.
    ORGANISM  unidentified
            unclassified.
  REFERENCE   1  (bases 1 to 50)
    AUTHORS   Rose-John,S.
    TITLE     CONJUGATE FOR MODIFYING INTERACTIONS BETWEEN PROTEINS
    JOURNAL   Patent: WO 9732891-A 9 12-SEP-1997;
            ANGEWANDTE GENTECHNOLOGIE SYST (DE); ROSE JOHN STEFAN (DE)
  FEATURES             Location/Qualifiers
       source          1. .50
                       /organism="unidentified"
                       /db_xref="taxon:32644"
  BASE COUNT       13 a     26 c      4 g      7 t
  ORIGIN

    Query Match             1.3%;  Score 22.4;  DB 6;  Length 50;
    Best Local Similarity   66.7%;  Pred. No. 1.6e+06;
    Matches   32;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

  Qy    958 CGGCAGAAGGTGCTACACCGAGACCTCAAGCCCCAGAACCTGCTCATC 1005
            || |||||  | |  | ||   ||||| | | |||||||||| | | ||
  Db      2 CGACAGAACCTCCACCTCCAGAACCTCCACCTCCAGAACCTCCACCTC 49


  RESULT 4
  I84671
  LOCUS       I84671                    47 bp    DNA     linear   PAT 04-APR-1998
  DEFINITION  Sequence 5 from patent US 5696086.
  ACCESSION   I84671
  VERSION     I84671.1  GI:3022191
  KEYWORDS    .
  SOURCE      Unknown.
    ORGANISM  Unknown.
            Unclassified.
  REFERENCE   1  (bases 1 to 47)
    AUTHORS   Avraham,H.Karsenty. and Godowski,P.J.
    TITLE     Methods and kits using macrophage stimulating protein
    JOURNAL   Patent: US 5696086-A 5 09-DEC-1997;
  FEATURES             Location/Qualifiers
       source          1. .47
                       /organism="unknown"
  BASE COUNT        8 a     16 c     12 g     11 t
  ORIGIN

    Query Match             1.3%;  Score 22.2;  DB 6;  Length 47;
    Best Local Similarity   69.8%;  Pred. No. 1.8e+06;
    Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

  Qy    241 GGCGGCAGTGACCCTGGAGAGGCCCCCACACGTGCTGCTCCTG 283
            ||  | |   ||| ||| | ||| ||||| | ||||||| |||
  Db      2 GGACGAATCCACCATGGGGTGGCTCCCACTCCTGCTGCTTCTG 44


  RESULT 5
  AX159452
  LOCUS       AX159452                  50 bp    DNA     linear   PAT 22-JUN-2001
  DEFINITION  Sequence 2780 from Patent WO0140521.
  ACCESSION   AX159452
  VERSION     AX159452.1  GI:14540783
  KEYWORDS    .
  SOURCE      human.
    ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
```

```
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Shimkets,R.A. and Leach,M.
  TITLE     Nucleic acids containing single nucleotide polymorphisms and
            methods of use thereof
  JOURNAL   Patent: WO 0140521-A 2780 07-JUN-2001;
            Curagen Corporation (US)
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
     misc_feature    25. .26
                     /note="Nucleotide deleted between bases 25 and 26
                     Accession number cg42460243"
     misc_feature    26
                     /note="2 of 2 allelic variants (2779 is other entry)"
BASE COUNT        7 a     22 c     13 g      8 t
ORIGIN

  Query Match             1.3%;  Score 22;  DB 6;  Length 50;
  Best Local Similarity   73.7%;  Pred. No. 2e+06;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1642 CGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          | ||| ||| | |||||||||||||| |  ||   ||||
Db     11 CTGCTTGAGCGCTGCCACACCCCTCTCGTGGGCCCCCC 48


RESULT 6
AX248673
LOCUS       AX248673                 31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 752 from Patent WO0166800.
ACCESSION   AX248673
VERSION     AX248673.1  GI:15863296
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 752 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        8 a     11 c      7 g      4 t      1 others
ORIGIN

  Query Match             1.2%;  Score 21.6;  DB 6;  Length 31;
  Best Local Similarity   80.0%;  Pred. No. 2.6e+06;
  Matches   24;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    979 GACCTCAAGCCCCAGAACCTGCTCATCAAC 1008
          ||| ||||||||||:||||||||  |  ||
Db      2 GACATCAAGCCCCAKAACCTGCTGGTGGAC 31


RESULT 7
AX182243/c
LOCUS       AX182243                 42 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 53 from Patent WO0142441.
ACCESSION   AX182243
VERSION     AX182243.1  GI:15133518
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1  (bases 1 to 42)
  AUTHORS   Reddy,S.I., Sadhu,L.I., Shukla,V.C. and Ferraiolo,G.I.
  TITLE     Plastid transformation
  JOURNAL   Patent: WO 0142441-A 53 14-JUN-2001;
            International Centre for Genetic Engineering and Biotechnology (IT)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT       15 a     10 c     11 g      6 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 6;  Length 42;
  Best Local Similarity   80.6%;  Pred. No. 2.8e+06;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    270 ACGTGCTGCTCCTGGGGAACTTCGTTCTGCA 300
          |||| | | |||||| || ||||| ||||||
Db     35 ACGTACGGGTCCTGGCGACCTTCGATCTGCA 5


RESULT 8
AX382049/c
LOCUS       AX382049                 42 bp    DNA     linear   PAT 18-MAR-2002
DEFINITION  Sequence 53 from Patent WO0206497.
ACCESSION   AX382049
VERSION     AX382049.1  GI:19576871
KEYWORDS    .
SOURCE      synthetic construct.
  ORGANISM  synthetic construct
            artificial sequences.
REFERENCE   1
  AUTHORS   Reddy,V.S. and Sadhu,L.
  TITLE     Transplastomic plants
  JOURNAL   Patent: WO 0206497-A 53 24-JAN-2002;
            International Centre for Genetic Engineering and Biotechnology (IT)
FEATURES             Location/Qualifiers
     source          1. .42
                     /organism="synthetic construct"
                     /db_xref="taxon:32630"
                     /note="Primer"
BASE COUNT       15 a     10 c     11 g      6 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 6;  Length 42;
  Best Local Similarity   80.6%;  Pred. No. 2.8e+06;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    270 ACGTGCTGCTCCTGGGGAACTTCGTTCTGCA 300
          |||| | | |||||| || ||||| ||||||
Db     35 ACGTACGGGTCCTGGCGACCTTCGATCTGCA 5


RESULT 9
AR032544
LOCUS       AR032544                 46 bp    DNA     linear   PAT 29-SEP-1999
DEFINITION  Sequence 156 from patent US 5869241.
ACCESSION   AR032544
VERSION     AR032544.1  GI:5948149
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Method of determining DNA sequence preference of a DNA-binding
            molecule
  JOURNAL   Patent: US 5869241-A 156 09-FEB-1999;
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unknown"
BASE COUNT        9 a     14 c     16 g      7 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 6;  Length 46;
```

```
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 10
AR209208
LOCUS       AR209208                  46 bp    DNA     linear   PAT 20-JUN-2002
DEFINITION  Sequence 156 from patent US 6384208.
ACCESSION   AR209208
VERSION     AR209208.1  GI:21510563
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence directed DNA binding molecules compositions and methods
  JOURNAL   Patent: US 6384208-A 156 07-MAY-2002;
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unknown"
BASE COUNT        9 a      14 c      16 g       7 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 6;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 11
I29284
LOCUS       I29284                    46 bp    DNA     linear   PAT 06-FEB-1997
DEFINITION  Sequence 156 from patent US 5578444.
ACCESSION   I29284
VERSION     I29284.1  GI:1820075
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M., Turin,L.M. and Fry,K.E.
  TITLE     Sequence-directed DNA-binding molecules compositions and methods
  JOURNAL   Patent: US 5578444-A 156 26-NOV-1996;
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unknown"
BASE COUNT        9 a      14 c      16 g       7 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 6;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 12
I90958
LOCUS       I90958                    46 bp    DNA     linear   PAT 01-DEC-1998
DEFINITION  Sequence 156 from patent US 5726014.
ACCESSION   I90958
VERSION     I90958.1  GI:3935428
KEYWORDS    .
SOURCE      Unknown.
  ORGANISM  Unknown.
            Unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Edwards,C.A., Cantor,C.R., Andrews,B.M. and Turin,L.M.
  TITLE     Screening assay for the detection of DNA-binding molecules
  JOURNAL   Patent: US 5726014-A 156 10-MAR-1998;
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unknown"
BASE COUNT        9 a      14 c      16 g       7 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 6;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 2.8e+06;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 13
AX248015
LOCUS       AX248015                  31 bp    DNA     linear   PAT 28-SEP-2001
DEFINITION  Sequence 94 from Patent WO0166800.
ACCESSION   AX248015
VERSION     AX248015.1  GI:15862638
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 31)
  AUTHORS   Cargill,M., Ireland,J.S. and Lander,E.S.
  TITLE     Human single nucleotide polymorphisms
  JOURNAL   Patent: WO 0166800-A 94 13-SEP-2001;
            WHITEHEAD INSTITUTE FOR BIOMEDICAL RESEARCH (US)
FEATURES             Location/Qualifiers
     source          1. .31
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT        5 a      9 c      8 g      8 t      1 others
ORIGIN

  Query Match             1.2%;  Score 21;  DB 6;  Length 31;
  Best Local Similarity   77.4%;  Pred. No. 3.5e+06;
  Matches   24;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    577 GTCAGCCTATCTGAGATTGGCTTTGGGAAAC 607
          | |  ||| || || :|||||||||||||||
Db      1 GCCTCCCTGTCAGACMTTGGCTTTGGGAAAC 31


RESULT 14
AX186238/c
LOCUS       AX186238                  46 bp    DNA     linear   PAT 06-AUG-2001
DEFINITION  Sequence 1933 from Patent WO0142467.
ACCESSION   AX186238
VERSION     AX186238.1  GI:15137666
KEYWORDS    .
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Schlegel,R., Deeds,J., Berger,A. and Zhao,X.
  TITLE     Genes, compositions, kits, and methods for identification,
            assessment, prevention, and therapy of cervical cancer
  JOURNAL   Patent: WO 0142467-A 1933 14-JUN-2001;
            Millennium Predictive Medicine, Inc. (US)
FEATURES             Location/Qualifiers
     source          1. .46
```

```
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
BASE COUNT         7 a      13 c      18 g       6 t       2 others
ORIGIN

  Query Match             1.2%;  Score 21;  DB 6;  Length 46;
  Best Local Similarity   71.1%;  Pred. No. 3.5e+06;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy     550 AAGCCCCTCAGCCGCCGCCTCCGTCGTGTCAGCCTATC 587
           ||||  ||| || |||||| |||  | ||   ||||||
Db      41 AAGCGTCTCTGCAGCCGCCNCCGCGGAGTGCTCCTATC 4


RESULT 15
A98791/c
LOCUS       A98791                    46 bp    DNA     linear   PAT 26-JAN-2000
DEFINITION  Sequence 24 from Patent WO9910358.
ACCESSION   A98791
VERSION     A98791.1  GI:6781812
KEYWORDS    .
SOURCE      unidentified.
  ORGANISM  unidentified
            unclassified.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Hegemann,P.
  TITLE     METHOD FOR PRODUCING NUCLEIC ACID POLYMERS
  JOURNAL   Patent: WO 9910358-A 24 04-MAR-1999;
            HEGEMANN PETER (DE)
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="unidentified"
                     /db_xref="taxon:32644"
BASE COUNT         9 a      13 c      16 g       8 t
ORIGIN

  Query Match             1.2%;  Score 20.8;  DB 6;  Length 46;
  Best Local Similarity   78.1%;  Pred. No. 3.9e+06;
  Matches   25;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy     500 TGCCTGAGGGCTACCTGGAGAAGCTGACCCTC 531
           |||| ||||||||| || || |||  ||| ||
Db      37 TGCCCGAGGGCTACGTGCAGGAGCGCACCATC 6


Search completed: March  4, 2003, 00:06:28
Job time : 4794 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 19:29:35 ; Search time 427 Seconds
(without alignments)
9203.131 Million cell updates/sec

Title: US-10-017-621-3
Perfect score: 1745
Sequence: 1 tggaagcagcgtaaaggatg.....gttcacctgcccacttgtcc 1745

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 2185239 seqs, 1125999159 residues

Total number of hits satisfying chosen parameters: 2166140

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : N_Geneseq_101002:*
1: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1980.DAT:*
2: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1981.DAT:*
3: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1982.DAT:*
4: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1983.DAT:*
5: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1984.DAT:*
6: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1985.DAT:*
7: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1986.DAT:*
8: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1987.DAT:*
9: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1988.DAT:*
10: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1989.DAT:*
11: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1990.DAT:*
12: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1991.DAT:*
13: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1992.DAT:*
14: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1993.DAT:*
15: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1994.DAT:*
16: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1995.DAT:*
17: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1996.DAT:*
18: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1997.DAT:*
19: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1998.DAT:*
20: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA1999.DAT:*
21: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2000.DAT:*
22: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001A.DAT:*
23: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2001B.DAT:*
24: /SIDS2/gcgdata/geneseq/geneseqn-embl/NA2002.DAT:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
c	1	24.8	1.4	50	22	AAL34335	Human SNP oligonuc
	2	22.4	1.3	33	24	ABA04099	Human Cdk5 related
c	3	22.4	1.3	33	24	ABA04100	Human Cdk5 related
	4	22	1.3	31	22	AAI30264	Human single nucle
	5	22	1.3	50	22	AAI75839	Human silent SNP c
	6	21.4	1.2	31	22	AAI29606	Human single nucle
c	7	21.4	1.2	42	22	AAH22523	PCR primer SR53 fo
c	8	21.4	1.2	42	24	AAD29563	ifnG coding region
	9	21.4	1.2	46	15	AAQ69406	Human H1 histone g
	10	21.4	1.2	46	18	AAT63868	Human H1 histone g
	11	21.4	1.2	46	20	AAX17156	Test sequence from
	12	21.4	1.2	46	24	ABK82647	DNA binding molecu
	13	21	1.2	21	22	AAH62195	PCTAIRE-1 polymorp
c	14	21	1.2	46	22	AAH70659	Human cervical can
c	15	21	1.2	50	22	AAL34286	Human SNP oligonuc
c	16	20.8	1.2	46	20	AAX22932	DE19736591 primer
	17	20.6	1.2	45	16	AAT07598	RT-PCR primer/prob
	18	20.6	1.2	45	16	AAT00670	Primer 143 for hum
c	19	20.6	1.2	50	22	AAL27875	Human SNP oligonuc
c	20	20.4	1.2	32	24	AAI72077	Xcds1 degenerate p
	21	20.4	1.2	48	14	AAQ50230	HIV pol INS mutage
c	22	20.2	1.2	36	19	AAV46356	PCR primer for ser
	23	20.2	1.2	40	16	AAQ76190	Primer for amplify
	24	20.2	1.2	40	18	AAT91033	Human 4-1BB 3' PCR
	25	20.2	1.2	40	24	ABL54053	Human cytokine rec
	26	20.2	1.2	41	24	ABK48869	Human proton-adeno
	27	20.2	1.2	41	24	ABK48870	Human proton-adeno
c	28	20.2	1.2	45	22	AAD17287	Human prostate spe
	29	20.2	1.2	49	20	AAZ31379	MUSIGHAEI Mouse Ig
	30	20.2	1.2	50	19	AAV59127	Reverse PCR primer
c	31	20.2	1.2	50	22	AAS43538	Corneodesmosin sin
	32	20	1.1	41	19	AAV37843	CD4+ human T-lymph
c	33	20	1.1	41	24	AAL43819	Human large protei
	34	20	1.1	46	21	AAZ91309	IL-6R and IL-6 fus
	35	20	1.1	47	21	AAZ67885	Human map-related
	36	20	1.1	50	22	AAL29783	Human SNP oligonuc
	37	20	1.1	50	22	AAL34645	Human SNP oligonuc
	38	19.8	1.1	46	21	AAC82589	Hammerhead ribozym
	39	19.8	1.1	48	18	AAT76056	Human A2b adenosin
	40	19.8	1.1	48	20	AAX53859	Human adenosine A2
	41	19.8	1.1	48	21	AAF19424	Human adenosine A2
	42	19.8	1.1	48	21	AAA33302	Low adenosine anti
	43	19.8	1.1	48	21	AAA03704	Human adenosine A1
c	44	19.8	1.1	48	22	AAH74232	Nucleotide sequenc
	45	19.6	1.1	31	22	AAI30029	Human single nucle

ALIGNMENTS

RESULT 1
AAL34335/c
ID AAL34335 standard; DNA; 50 BP.
XX
AC AAL34335;
XX
DT 24-JAN-2002 (first entry)
XX
DE Human SNP oligonucleotide #7543.
XX
KW Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW complement related protein; cytochrome; kinesin; cytokine; interferon;
KW interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW multifactorial disease; autoimmune disease; infection;
KW nervous system disease; ss.
XX
OS Homo sapiens.
XX
PN WO200147944-A2.
XX
PD 05-JUL-2001.
XX
PF 28-DEC-2000; 2000WO-US35498.
XX
PR 28-DEC-1999; 99US-0173419.
PR 27-DEC-2000; 2000US-0173419.
XX
PA (CURA-) CURAGEN CORP.
XX

PI Shimkets RA, Leach M;
XX
DR WPI; 2001-465210/50.
XX
PT Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT oncogenes and histones, useful for diagnosing and treating, e.g.
PT cancer, autoimmune diseases and infections -
XX
PS Claim 1; Page 3563; 4143pp; English.
XX
CC The present invention relates to oligonucleotides encoding polymorphic
CC variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC histones, kinases, colony stimulating factors, complement related
CC proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC G-protein coupled receptors and thioesterases. The present sequence is
CC one such oligonucleotide. The oligonucleotides and the peptides encoded
CC by them may be used in the prevention, diagnosis and treatment of
CC diseases associated with inappropriate expression of the proteins listed
CC above. Disorders that may be prevented, diagnosed and/or treated include
CC multifactorial diseases with a genetic component, such as autoimmune
CC diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC leukaemia), diseases of the nervous system and an infection of pathogenic
CC organisms.
XX
SQ Sequence 50 BP; 7 A; 13 C; 14 G; 16 T; 0 other;

Query Match 1.4%; Score 24.8; DB 22; Length 50;
Best Local Similarity 72.7%; Pred. No. 7.8e+03;
Matches 32; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

Qy 1619 CAGACCGAGGCCCCAGCAGGCAGCGGCTGGAGGGATGCCACACC 1662
||||||||| |||||||| | ||||||| | |||||
Db 45 CAGACCGAGGACCCAGCAGACTCATGCTGGAGAATTATGACACC 2

RESULT 2
ABA04099
ID ABA04099 standard; DNA; 33 BP.
XX
AC ABA04099;
XX
DT 21-FEB-2002 (first entry)
XX
DE Human Cdk5 related PCR primer SEQ ID NO:18.
XX
KW Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;
KW amyloid precursor protein; APP; Cdk5; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200182967-A1.
XX
PD 08-NOV-2001.
XX
PF 25-APR-2001; 2001WO-JP03555.
XX
PR 28-APR-2000; 2000JP-0131037.
XX
PA (YAMA) YAMANOUCHI PHARM CO LTD.
PA (SUZU/) SUZUKI T.
XX
PI Suzuki T, Watanabe T, Kawabata S, Hachiya S;
XX
DR WPI; 2002-026209/03.
XX
PT Medicinal compositions for the treatment of dementia and Alzheimer's
PT disease, comprise compounds that suppress beta amyloid production -
XX
PS Example 6; Page 23; 62pp; Japanese.
XX
CC The present invention describes medicinal compositions (I) inhibiting
CC beta-amyloid production comprising an active component a substance that
CC inhibits the activity of cyclin-dependent kinase (CDK). Also described
CC are: (1) a method for screening compounds for their ability to inhibit
CC the production of beta-amyloid by contacting with beta-amyloid producing
CC cells; and (2) screening kits. (I) have nootropic and neuroprotective
CC activities. (I) suppress the phosphorylation of amyloid precursor protein
CC (APP) which is an essential step in the production of beta-amyloid. (I)
CC can be used in the treatment and prevention of neurodegenerative diseases
CC such as dementia and Alzheimer's disease. The present sequence represents
CC a PCR primer which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 33 BP; 6 A; 6 C; 11 G; 10 T; 0 other;

Query Match 1.3%; Score 22.4; DB 24; Length 33;
Best Local Similarity 81.2%; Pred. No. 2.8e+04;
Matches 26; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

Qy 1018 GAGCTCAAGCTGGCTGACTTTGGCCTGGCCCG 1049
||||| || ||||| | ||||||||||| ||
Db 2 GAGCTGAAATTGGCTAATTTTGGCCTGGCTCG 33

RESULT 3
ABA04100/c
ID ABA04100 standard; DNA; 33 BP.
XX
AC ABA04100;
XX
DT 21-FEB-2002 (first entry)
XX
DE Human Cdk5 related PCR primer SEQ ID NO:19.
XX
KW Human; beta-amyloid; cyclin-dependent kinase inhibitor; nerve cell;
KW amyloid precursor protein; APP; Cdk5; PCR primer; ss.
XX
OS Homo sapiens.
XX
PN WO200182967-A1.
XX
PD 08-NOV-2001.
XX
PF 25-APR-2001; 2001WO-JP03555.
XX
PR 28-APR-2000; 2000JP-0131037.
XX
PA (YAMA) YAMANOUCHI PHARM CO LTD.
PA (SUZU/) SUZUKI T.
XX
PI Suzuki T, Watanabe T, Kawabata S, Hachiya S;
XX
DR WPI; 2002-026209/03.
XX
PT Medicinal compositions for the treatment of dementia and Alzheimer's
PT disease, comprise compounds that suppress beta amyloid production -
XX
PS Example 6; Page 23; 62pp; Japanese.
XX
CC The present invention describes medicinal compositions (I) inhibiting
CC beta-amyloid production comprising an active component a substance that
CC inhibits the activity of cyclin-dependent kinase (CDK). Also described
CC are: (1) a method for screening compounds for their ability to inhibit
CC the production of beta-amyloid by contacting with beta-amyloid producing
CC cells; and (2) screening kits. (I) have nootropic and neuroprotective
CC activities. (I) suppress the phosphorylation of amyloid precursor protein
CC (APP) which is an essential step in the production of beta-amyloid. (I)
CC can be used in the treatment and prevention of neurodegenerative diseases
CC such as dementia and Alzheimer's disease. The present sequence represents
CC a PCR primer which is used in the exemplification of the present
CC invention.
XX
SQ Sequence 33 BP; 10 A; 11 C; 6 G; 6 T; 0 other;

```
  Query Match             1.3%;  Score 22.4;  DB 24;  Length 33;
  Best Local Similarity   81.2%;  Pred. No. 2.8e+04;
  Matches   26;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy   1018 GAGCTCAAGCTGGCTGACTTTGGCCTGGCCCG 1049
          ||||| ||  ||||| | ||||||||||| ||
Db     32 GAGCTGAAATTGGCTAATTTTGGCCTGGCTCG 1


RESULT 4
AAI30264
ID   AAI30264 standard; DNA; 31 BP.
XX
AC   AAI30264;
XX
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) 97.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,T)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 75; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 8 A; 11 C; 8 G; 4 T; 0 other;

  Query Match             1.3%;  Score 22;  DB 22;  Length 31;
  Best Local Similarity   83.3%;  Pred. No. 3.4e+04;
  Matches   25;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    979 GACCTCAAGCCCCAGAACCTGCTCATCAAC 1008
          ||| ||||||||||||||||||  |  ||
Db      2 GACATCAAGCCCCAGAACCTGCTGGTGGAC 31


RESULT 5
AAI75839
ID   AAI75839 standard; DNA; 50 BP.
XX
AC   AAI75839;
XX
DT   09-NOV-2001  (first entry)
XX
DE   Human silent SNP containing nucleic acid SEQ:2780.
XX
KW   Human; single nucleotide polymorphism; SNP; genome; gene therapy;
KW   protein therapy; vaccine; probe; diagnostic assay; detection;
KW   quantitation; restorative therapy; polymorphic; ds.
XX
OS   Homo sapiens.
XX
PN   WO200140521-A2.
XX
PD   07-JUN-2001.
XX
PF   30-NOV-2000; 2000WO-US32758.
XX
PR   30-NOV-1999;   99US-0168138.
PR   29-NOV-2000; 2000US-0726173.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-356160/37.
XX
PT   Polymorphic nucleic acid sequences, useful in genetic testing and
PT   therapy -
XX
PS   Claim 1; Page 901; 2653pp; English.
XX
CC   AAI73060 to AAI79867 represent isolated human polymorphic polynucleotide
CC   sequences (I), which contain single nucleotide polymorphisms (SNPs).
CC   AAM53114 to AAM53329 represent peptides related to human polymorphic
CC   polynucleotide sequences. The sequences can be used in gene and protein
CC   therapy, and in vaccine production. (I) and the polypeptides encoded by
CC   them may be used in the prevention, diagnosis and treatment of diseases
CC   associated with inappropriate expression of polymorphic polypeptides.
CC   For example, (I) may be used to treat disorders by rectifying mutations
CC   or deletions in a patient's genome that affect the activity of
CC   polypeptides by expressing inactive proteins or to supplement the
CC   patients own production of polypeptide. Additionally, (I) and its
CC   complementary sequences may also be used as DNA probes in diagnostic
CC   assays to detect and quantitate the presence of similar nucleic acids
CC   in samples, and therefore which patients may be in need of restorative
CC   therapy. The polypeptides encoded by (I) may be used as antigens in the
CC   production of antibodies specific for polymorphic polypeptides. The
CC   antibodies may also be used to down regulate expression and activity.
CC   The antibodies may also be used as diagnostic agents for detecting the
CC   presence of polymorphic polypeptides in samples.
XX
SQ   Sequence 50 BP; 7 A; 22 C; 13 G; 8 T; 0 other;

  Query Match             1.3%;  Score 22;  DB 22;  Length 50;
  Best Local Similarity   73.7%;  Pred. No. 4.2e+04;
  Matches   28;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1642 CGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          | ||| ||| | ||||||||||||| |  ||   ||||
Db     11 CTGCTTGAGCGCTGCCACACCCCTCTCGTGGGCCCCCC 48


RESULT 6
AAI29606
ID   AAI29606 standard; DNA; 31 BP.
XX
AC   AAI29606;
XX
```

```
DT   18-OCT-2001  (first entry)
XX
DE   Human single nucleotide polymorphism (SNP) PCTAIRE3 1.
XX
KW   Human; resequence; genotype; disease; forensic; paternity testing;
KW   single nucleotide polymorphism; SNP; ss.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(16,C)
FT                   /*tag=  a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200166800-A2.
XX
PD   13-SEP-2001.
XX
PF   07-MAR-2001; 2001WO-US07268.
XX
PR   07-MAR-2000; 2000US-0187510.
PR   22-MAY-2000; 2000US-0206129.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-522952/57.
XX
PT   Nucleic acid molecules from the human genome which include polymorphic
PT   sites, useful in methods for predicting the presence, absence or
PT   severity of a particular phenotype or disorder (e.g. diabetes)
PT   associated with a particular genotype -
XX
PS   Claim 1; Page 34; 145pp; English.
XX
CC   The invention relates to the identification of nucleic acid molecules
CC   (AAI29513-AAI31314) from the human genome which include polymorphic sites
CC   which can predispose individuals to disease. Various genes from a number
CC   of individuals were resequenced and single nucleotide polymorphisms
CC   (SNPs) in these genes discovered. The method is useful for predicting the
CC   presence, absence or severity of a particular phenotype or disorder (e.g.
CC   diabetes) associated with a particular genotype. The nucleic acids
CC   containing the polymorphic sites may be useful in forensics and paternity
CC   testing.
XX
SQ   Sequence 31 BP; 6 A; 9 C; 8 G; 8 T; 0 other;

  Query Match             1.2%;  Score 21.4;  DB 22;  Length 31;
  Best Local Similarity   80.6%;  Pred. No. 4.9e+04;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    577 GTCAGCCTATCTGAGATTGGCTTTGGGAAAC 607
          | |  ||| || || ||||||||||||||||
Db      1 GCCTCCCTGTCAGACATTGGCTTTGGGAAAC 31


RESULT 7
AAH22523/c
ID   AAH22523 standard; DNA; 42 BP.
XX
AC   AAH22523;
XX
DT   22-AUG-2001  (first entry)
XX
DE   PCR primer SR53 for amplifying a ifnG coding region.
XX
KW   Transplastome; plastome; plastid; chloroplast; transgene; plant;
KW   ifnG; PCR primer; ss.
XX
OS   Synthetic.
XX
PN   WO200142441-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-EP12446.
XX
PR   08-DEC-1999;   99GB-0029075.
PR   14-JUL-2000; 2000GB-0017369.
XX
PA   (ITGE-) INT CENT GENETIC ENG & BIOTECHNOLOGY.
XX
PI   Reddy S,  Sadhu L,  Shukla V,  Ferraiolo G;
XX
DR   WPI; 2001-381671/40.
XX
PT   Obtaining a stable transplastome for producing a  transplastomic cell,
PT   plant or seed, comprises transforming a recipient plastome with a
PT   polynucleotide comprising a 5' and 3' sequence homologous to the
PT   recipient -
XX
PS   Example 12; Page 127; 128pp; English.
XX
CC   The invention relates to a method of obtaining a stable transplastome,
CC   by transforming a recipient plastome (RP) with a polynucleotide having a
CC   5' sequence homologous to a region of RP, and joined to it, a sequence
CC   heterologous to RP comprising a coding region operably linked to
CC   regulatory region capable of securing expression of coding region in the
CC   plastid and joined to it, and a 3' sequence homologous to a region of RP.
CC   The method is useful for obtaining a transplastomic plastid, by
CC   transforming a plastome within a plastid such as proplastid, amyloplast,
CC   chromoplast, etioplast or leucoplast, preferably chloroplast. The method
CC   is useful for obtaining a transplastomically expressed protein. The
CC   method provides high, uniform, reliable expression of transgenes in
CC   plants, with stable inheritance of the trait by avoiding the potential
CC   for the dangerous spread of transgenes to the ecosystem. The present
CC   sequence represents a PCR primer for amplifying a ifnG coding region,
CC   used in generating expression vectors for ifnG in chloroplasts.
XX
SQ   Sequence 42 BP; 15 A; 10 C; 11 G; 6 T; 0 other;

  Query Match             1.2%;  Score 21.4;  DB 22;  Length 42;
  Best Local Similarity   80.6%;  Pred. No. 5.6e+04;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    270 ACGTGCTGCTCCTGGGGAACTTCGTTCTGCA 300
          |||| | | |||||| || ||||| |||||||
Db     35 ACGTACGGGTCCTGGCGACCTTCGATCTGCA 5


RESULT 8
AAD29563/c
ID   AAD29563 standard; DNA; 42 BP.
XX
AC   AAD29563;
XX
DT   07-MAY-2002  (first entry)
XX
DE   ifnG coding region DNA amplifying PCR primer, SR53.
XX
KW   Transgenic plant; transplastomic plant; medicament; PCR primer; ss.
XX
OS   Unidentified.
XX
PN   WO200206497-A2.
XX
PD   24-JAN-2002.
XX
PF   13-JUL-2001; 2001WO-EP08132.
XX
PR   14-JUL-2000; 2000GB-0017397.
XX
PA   (ITGE-) INT CENT GENETIC ENG & BIOTECHNOLOGY.
XX
PI   Reddy VS,  Sadhu L;
```

XX
DR WPI; 2002-171810/22.
XX
PT Producing a protein of interest, e.g., a pharmaceutically active
PT protein, comprises expressing a polynucleotide fusion construct in a
PT plastid and producing a fusion protein comprising the protein of
PT interest -
XX
PS Example 1; Page 90; 92pp; English.
XX
CC The patent discloses a method of producing a protein of interest which
CC involves expressing a polynucleotide fusion construct in a plastid to
CC produce a fusion protein comprising the protein of interest where the
CC construct comprises a polynucleotide coding sequence of the protein of
CC interest operably linked to a polynucleotide coding sequence of a fusion
CC protein partner. The methods of the invention are useful for producing a
CC protein of interest which comprises a human protein or its biologically
CC active variant or fragment, a pharmaceutically active protein, an IFN
CC (interferon), its biologically active variant or fragment, a human IFN-
CC gamma or its biologically active variant or fragment. They are useful
CC for the production of transgenic plants. Methods of the invention are
CC also useful for the generation of transplastomic plant cells, plants
CC and seeds. The protein of interest obtained by the methods of the
CC invention is useful for the manufacture of a medicament for treating
CC a disease condition. The present DNA sequence is a PCR primer which
CC is used for amplifying ifnG coding region. This primer is used in
CC the exemplification of the invention to generate pGUSIFNG expression
CC vector for ifnG in chloroplasts.
XX
SQ Sequence 42 BP; 15 A; 10 C; 11 G; 6 T; 0 other;

Query Match 1.2%; Score 21.4; DB 24; Length 42;
Best Local Similarity 80.6%; Pred. No. 5.6e+04;
Matches 25; Conservative 0; Mismatches 6; Indels 0; Gaps 0;

```
Qy    270 ACGTGCTGCTCCTGGGGAACTTCGTTCTGCA 300
          |||| | | |||||| || ||||| ||||||
Db     35 ACGTACGGGTCCTGGCGACCTTCGATCTGCA 5
```

RESULT 9
AAQ69406
ID AAQ69406 standard; DNA; 46 BP.
XX
AC AAQ69406;
XX
DT 27-FEB-1995 (first entry)
XX
DE Human H1 histone gene FNC16, target region.
XX
KW DNA protein-binding assay; test sequence; screening sequence;
KW promoter; target; TATA box; Herpes Simplex Virus; HSV;
KW origin of replication; UL9; transcription factor; TFIID: ds.
XX
OS Synthetic.
XX
PN WO9414980-A.
XX
PD 07-JUL-1994.
XX
PF 20-DEC-1993; 93WO-US12388.
XX
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1994-234711/28.
XX
PT Sequence-directed DNA-binding molecules - useful in
PT pharmaceuticals and as molecular reagents
XX
PS Claim 28; Page 290; 587pp; English.
XX
CC A DNA protein-binding assay is provided, useful for screening
CC libraries of synthetic or biological cpds. for their ability
CC to bind DNA test sequences. The assay is versatile in that any
CC number of test sequences can be tested by placing the test sequence
CC adjacent to a defined protein-binding screening sequence. Binding
CC of mols. to these test sequences changes the binding characteristics
CC of the protein mol. to its cognate binding sequence. When such a mol.
CC binds the test sequence, the equilibrium of the DNA:protein complexes
CC is disturbed, generating changes in the concentration of free DNA probe.
CC One application of this method is to eucaryotic general transcription
CC factors (e.g. TFIID), where the target region is typically selected
CC from DNA sequences adjacent to the binding site for the eucaryotic
CC transcription factor. Numerous exemplary test sequences are given:
CC the sequences in AAQ69251-731 and AAQ69850 correspond to promoter
CC targets (typically, TATA box-contg. sites) for human genes and the
CC sequences in AAQ69732-849 correspond to promoter targets for viral genes.
CC The test sequences may also be randomly generated. DNA:protein
CC interaction may be used for screening purposes, e.g. the Herpes Simplex
CC Virus (HSV) origin of replication and UL9 (see AAQ69851-52, AAQ69865 and
CC AAQ69891).
XX
SQ Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.2%; Score 21.4; DB 15; Length 46;
Best Local Similarity 71.8%; Pred. No. 5.8e+04;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

```
Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  |||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40
```

RESULT 10
AAT63868
ID AAT63868 standard; DNA; 46 BP.
XX
AC AAT63868;
XX
DT 14-MAR-1997 (first entry)
XX
DE Human H1 histone gene FNC16 gene TFIID binding site.
XX
KW Duplex DNA; target region; binding characteristic; DNA binding protein;
KW TFIID; transcription factor; binding site; inhibition; enhance;
KW cancer; inherited genetic disorder; alpha-D-galactosidase A; ds.
XX
OS Homo sapiens.
XX
PN US5578444-A.
XX
PD 26-NOV-1996.
XX
PF 27-JUN-1991; 91US-0723618.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1997-020402/02.
XX
PT Altering binding characteristics of DNA binding proteins to duplex
PT DNA - by attaching specific small cpd. to target region close to the
PT protein's binding site, useful in treatment of viral disease, cancer
PT etc
XX

PS Claim 6; Column 177-178; 264pp; English.
XX
CC The sequences given in AAT63713-4312 represent duplex DNA's which act
CC as target regions in the method of the invention. The method for
CC altering the binding characteristics of a DNA-binding protein to duplex
CC DNA comprises contacting the duplex DNA with a small molecule which
CC binds sequence-specifically to a target region, where, when the small
CC molecule is bound to the target region, it is adjacent to, but not
CC overlapping by more than 4 bp, a binding site for a DNA-binding protein.
CC The small molecule is added at a concentration effective to alter the
CC binding of the DNA binding protein, pref. TFIID, to its binding site on
CC the duplex DNA. The binding of the small molecule may inhibit or
CC enhance the binding of the DNA-binding protein to its binding site. The
CC compounds isolated using this method are potentially useful as
CC therapeutic agents for treatment of any disease which involves a
CC specific DNA sequence, e.g. cancer, or inherited genetic disorders etc.
CC The method is suitable for screening large biological or chemical
CC libraries and allows determination of sequence-specific and relative
CC affinities of known DNA-binding agents for different DNA sequences.
CC The design of these duplex DNA's allows a single DNA:protein interaction
CC to be used for screening sequence-specific, or preferential, DNA binding
CC proteins that recognise almost any possible sequence (see also AAT49539-
CC 74).
XX
SQ Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.2%; Score 21.4; DB 18; Length 46;
Best Local Similarity 71.8%; Pred. No. 5.8e+04;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
|||| | ||| || |||| |||||||||||| ||
Db 2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40

RESULT 11
AAX17156
ID AAX17156 standard; DNA; 46 BP.
XX
AC AAX17156;
XX
DT 06-MAY-1999 (first entry)
XX
DE Test sequence from human H1 histone gene FNC16.
XX
KW Test sequence; DNA-binding molecule; screening sequence; human;
KW nucleic acid amplification; target; viral; ds.
XX
OS Homo sapiens.
XX
PN US5869241-A.
XX
PD 09-FEB-1999.
XX
PF 07-JUN-1995; 95US-0475228.
XX
PR 20-DEC-1993; 93US-0171389.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
PR 07-JUN-1995; 95US-0475228.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Andrews BM, Cantor CR, Edwards CA, Fry KE, Turin LM;
XX
DR WPI; 1999-152755/13.
XX
PT Determination of DNA sequence preference of a DNA-binding molecule -
PT based on inhibition of binding of protein to oligonucleotide
PT sequence attached to test sequence
XX
PS Claim 3; Columns 179-180; 270pp; English.
XX
CC Sequences AAX17001 to AAX17600 represent specifically claimed target
CC test sequences that are used in the method of the invention of
CC determining the DNA sequence preference of a DNA-binding molecule. The
CC method comprises: (i) adding a test molecule and a DNA-binding protein to
CC a mixture of duplex DNA test oligonucleotides, each of the test
CC oligonucleotides having a test sequence adjacent to a screening sequence,
CC where the screening sequence binds to the DNA-binding protein with a
CC binding affinity that is independent of the DNA sequence of the test
CC sequence, and where the mixture of duplex DNA test oligonucleotides
CC includes several test sequences; (ii) incubating the test molecule, the
CC mixture of duplex DNA test oligonucleotides and the DNA-binding protein
CC for a time sufficient to permit binding of the test molecule to test
CC sequences in the duplex DNA; (iii) separating unbound test
CC oligonucleotides from test oligonucleotides bound to binding protein;
CC (iv) amplifying the unbound test oligonucleotides; (v) repeating steps
CC (ii) to (iv); (vi) isolating the amplified test oligonucleotides; and
CC (vii) sequencing the isolated test oligonucleotides. Test sequences
CC AAX17001-X17481 and AAX17600 correspond to promoter targets for human
CC genes and test sequences AAX17482-X17599 correspond to promoter targets
CC for viral genes.
XX
SQ Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

Query Match 1.2%; Score 21.4; DB 20; Length 46;
Best Local Similarity 71.8%; Pred. No. 5.8e+04;
Matches 28; Conservative 0; Mismatches 11; Indels 0; Gaps 0;

Qy 1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
|||| | ||| || |||| |||||||||||| ||
Db 2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40

RESULT 12
ABK82647
ID ABK82647 standard; DNA; 46 BP.
XX
AC ABK82647;
XX
DT 27-AUG-2002 (first entry)
XX
DE DNA binding molecule screening method test sequence #156.
XX
KW DNA binding molecule screening; inhibition of transcription;
KW infection; human immunodeficiency virus; HIV; parasite; cancer;
KW cardiovascular; respiratory; gastrointestinal; endocrine; metabolic;
KW rheumatic; immunological; haematological; neurological;
KW psychiatric; dermatological; ophthalmological; musculo-skeletal;
KW urogenital disorder; ss.
XX
OS Synthetic.
XX
PN US6384208-B1.
XX
PD 07-MAY-2002.
XX
PF 15-JUL-1999; 99US-0354947.
XX
PR 20-DEC-1993; 93US-0171389.
PR 07-JUN-1995; 95US-0482080.
PR 27-JUN-1991; 91US-0723618.
PR 23-DEC-1992; 92US-0996783.
PR 17-SEP-1993; 93US-0123936.
XX
PA (GENE-) GENELABS TECHNOLOGIES INC.
XX
PI Edwards CA, Cantor CR, Andrews BM, Turin LM, Fry KE;
XX
DR WPI; 2002-442819/47.
XX
PT Decreasing transcriptional activity of genes for treating infections or
PT cancer, by administration of an agent that binds to two non-overlapping
PT regions of the gene -

```
XX
PS   Example 15; SEQ ID No 156; 98pp; English.
XX
CC   The invention relates to a method of decreasing transcriptional activity
CC   in a duplex deoxyribonucleic acid (DNA) template (T1) comprising
CC   contacting (T1) with a binding agent comprising at least one small duplex
CC   DNA-binding molecule (T2) coupled to at least one other small duplex-
CC   binding molecule that binds to a non-overlapping region of target
CC   sequence (TS). The method is useful for inhibiting transcription of a
CC   range of disease-related genes for treating infections (by viruses,
CC   including human immunodeficiency virus, bacteria, fungi, protozoa
CC   and parasites), cancer, cardiovascular, respiratory, gastrointestinal,
CC   endocrine/metabolic, rheumatic/immunological, haematological,
CC   neurological, psychiatric, dermatological, ophthalmological,
CC   musculo-skeletal, genetic or urogenital disorders. The method provides
CC   sequence-specific inhibition of transcription of pathological genes
CC   without affecting transcription of cellular genes regulated by the same
CC   transcription factor, and can be applied to regulation of any gene.
CC   ABK82492-ABK83155 represent DNA binding molecule test sequences used in
CC   the method of the invention.
XX
SQ   Sequence 46 BP; 9 A; 14 C; 16 G; 7 T; 0 other;

  Query Match             1.2%;  Score 21.4;  DB 24;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 5.8e+04;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
           ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 13
AAH62195
ID   AAH62195 standard; DNA; 21 BP.
XX
AC   AAH62195;
XX
DT   12-SEP-2001  (first entry)
XX
DE   PCTAIRE-1 polymorphism containing DNA fragment #96.
XX
KW   Single nucleotide polymorphism; SNP; human; cancer; inflammation;
KW   heart disease; paternity testing; forensic science; ds.
XX
OS   Homo sapiens.
XX
FH   Key             Location/Qualifiers
FT   Variation       replace(11,G)
FT                   /*tag= a
FT                   /standard_name= "single nucleotide polymorphism"
XX
PN   WO200138576-A2.
XX
PD   31-MAY-2001.
XX
PF   17-NOV-2000; 2000WO-US31639.
XX
PR   24-NOV-1999;   99US-0167334.
XX
PA   (WHED ) WHITEHEAD INST BIOMEDICAL RES.
XX
PI   Cargill M,  Ireland JS,  Lander ES;
XX
DR   WPI; 2001-367705/38.
XX
PT   New nucleic acid segments of the human genome, particularly from genes
PT   including polymorphic sites,for phenotype correlation, forensics,
PT   paternity testing, medicine and genetic analysis  -
XX
PS   Claim 1; Page 37; 80pp; English.
XX
CC   DNA sequences AAH62100 - AAH62688 represent segments of human genes which
CC   contain single nucleotide polymorphisms (SNPs). A method is included in
CC   the invention for analysing a nucleic acid sample, which consists of
CC   determining the base occupying any one of the polymorphic sites given in
CC   the SNP containing sequences. The nucleotide sequences can be used in the
CC   diagnosis or monitoring of diseases, such as cancer, inflammation, heart
CC   diseases, diseases of the cardiovascular system, and infection by
CC   microorganisms. The oligonucleotides are also useful in the manufacture
CC   of a medicament for the treatment or prophylaxis of the diseases, and as
CC   a pharmaceutical. SNP containing oligonucleotides are useful in
CC   applications such as phenotype correlation, forensics, paternity testing,
CC   medicine and genetic analysis.
XX
SQ   Sequence 21 BP; 9 A; 4 C; 6 G; 2 T; 0 other;

  Query Match             1.2%;  Score 21;  DB 22;  Length 21;
  Best Local Similarity   100.0%;  Pred. No. 5.4e+04;
  Matches   21;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy    702 CAAGGAGATCAGACTGGAACA 722
          |||||||||||||||||||||
Db      1 CAAGGAGATCAGACTGGAACA 21


RESULT 14
AAH70659/c
ID   AAH70659 standard; cDNA; 46 BP.
XX
AC   AAH70659;
XX
DT   19-SEP-2001  (first entry)
XX
DE   Human cervical cancer marker nucleic acid 1933.
XX
KW   Cervical cancer; cytostatic; pre-malignant condition; gene therapy; ss.
XX
OS   Homo sapiens.
XX
PN   WO200142467-A2.
XX
PD   14-JUN-2001.
XX
PF   08-DEC-2000; 2000WO-US33312.
XX
PR   08-DEC-1999;   99US-0169681.
PR   21-DEC-1999;   99US-0171350.
PR   14-MAR-2000; 2000US-0189315.
PR   12-MAY-2000; 2000US-0203791.
PR   09-JUN-2000; 2000US-0210600.
PR   21-JUL-2000; 2000US-0220114.
XX
PA   (MILL-) MILLENNIUM PREDICTIVE MEDICINE INC.
XX
PI   Schlegel R,  Deeds J,  Berger A,  Zhao X;
XX
DR   WPI; 2001-375006/39.
XX
PT   New isolated nucleic acid for diagnosing and treating cervical cancer
PT   and for assessing and detecting compounds for treating the cancer -
XX
PS   Claim 1; Page 415; 1051pp; English.
XX
CC   The invention relates to novel genes (AAH68727-AAH73383) associated with
CC   cervical cancer with cytostatic activity. The nucleic acids and encoded
CC   polypeptides are useful: to assess if a patient is afflicted with
CC   cervical cancer or has a pre-malignant condition; to monitor the
CC   progression of cervical cancer or a premalignant condition in a patient;
CC   and to select and/or assess the efficacy of a compound or therapy for
CC   inhibiting cervical cancer in a patient. The nucleic acids may also be
CC   useful for gene therapy.
XX
SQ   Sequence 46 BP; 7 A; 13 C; 18 G; 6 T; 2 other;

  Query Match             1.2%;  Score 21;  DB 22;  Length 46;
```

```
  Best Local Similarity   71.1%;  Pred. No. 7.3e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    550 AAGCCCCTCAGCCGCCGCCTCCGTCGTGTCAGCCTATC 587
          ||||  ||| || |||||| |||  | ||   ||||||
Db     41 AAGCGTCTCTGCAGCCGCCNCCGCGGAGTGCTCCTATC 4


RESULT 15
AAL34286/c
ID   AAL34286 standard; DNA; 50 BP.
XX
AC   AAL34286;
XX
DT   24-JAN-2002  (first entry)
XX
DE   Human SNP oligonucleotide #7494.
XX
KW   Immunosuppressive; immunostimulatory; antiinflammatory; cytostatic;
KW   neuroprotective; antimicrobial; gene therapy; vaccine; amylase; cancer;
KW   amyloid protein; angiopoietin; apoptosis related protein; cadherin;
KW   cyclin; polymerase; oncogene; histone; kinase; colony stimulating factor;
KW   complement related protein; cytochrome; kinesin; cytokine; interferon;
KW   interleukin; G-protein coupled receptor; thioesterase; inflammation;
KW   multifactorial disease; autoimmune disease; infection;
KW   nervous system disease; ss.
XX
OS   Homo sapiens.
XX
PN   WO200147944-A2.
XX
PD   05-JUL-2001.
XX
PF   28-DEC-2000; 2000WO-US35498.
XX
PR   28-DEC-1999;   99US-0173419.
PR   27-DEC-2000; 2000US-0173419.
XX
PA   (CURA-) CURAGEN CORP.
XX
PI   Shimkets RA,  Leach M;
XX
DR   WPI; 2001-465210/50.
XX
PT   Polymorphic nucleic acids encoding e.g. amylases, cyclins, polymerases,
PT   oncogenes and histones, useful for diagnosing and treating, e.g.
PT   cancer, autoimmune diseases and infections -
XX
PS   Claim 1; Page 3547; 4143pp; English.
XX
CC   The present invention relates to oligonucleotides encoding polymorphic
CC   variants of proteins related to amylases, amyloid proteins, angiopoietin,
CC   apoptosis related proteins, cadherin, cyclin, polymerase, oncogenes,
CC   histones, kinases, colony stimulating factors, complement related
CC   proteins, cytochromes, kinesins, cytokines, interferons, interleukins,
CC   G-protein coupled receptors and thioesterases. The present sequence is
CC   one such oligonucleotide. The oligonucleotides and the peptides encoded
CC   by them may be used in the prevention, diagnosis and treatment of
CC   diseases associated with inappropriate expression of the proteins listed
CC   above. Disorders that may be prevented, diagnosed and/or treated include
CC   multifactorial diseases with a genetic component, such as autoimmune
CC   diseases (e.g. rheumatoid arthritis, multiple sclerosis, diabetes,
CC   systemic lupus erythromatosus and Grave's disease), inflammation, cancer
CC   (e.g. cancers of the bladder, brain, breast, colon and kidney,
CC   leukaemia), diseases of the nervous system and an infection of pathogenic
CC   organisms.
XX
SQ   Sequence 50 BP; 1 A; 27 C; 11 G; 11 T; 0 other;

  Query Match             1.2%;  Score 21;  DB 22;  Length 50;
  Best Local Similarity   66.7%;  Pred. No. 7.6e+04;
  Matches   30;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    208 GAGCAGATAGGCCTGGATGAGAGTGGTGGTGGTGGCGGCAGTGAC 252
          ||||||| |||   ||||| | |    || || |  ||||| |||
Db     50 GAGCAGAGAGGAGCGGATGGGCGGCCGGGCGGCGAGGGCAGCGAC 6


Search completed: March  3, 2003, 22:46:23
Job time : 431 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 22:35:55 ; Search time 76 Seconds
(without alignments)
7041.460 Million cell updates/sec

Title: US-10-017-621-3
Perfect score: 1745
Sequence: 1 tggaagcagcgtaaaggatg.....gttcacctgcccacttgtcc 1745

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 441362 seqs, 153338381 residues

Total number of hits satisfying chosen parameters: 609818

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Issued_Patents_NA:*
1: /cgn2_6/ptodata/1/ina/5A_COMB.seq:*
2: /cgn2_6/ptodata/1/ina/5B_COMB.seq:*
3: /cgn2_6/ptodata/1/ina/6A_COMB.seq:*
4: /cgn2_6/ptodata/1/ina/6B_COMB.seq:*
5: /cgn2_6/ptodata/1/ina/PCTUS_COMB.seq:*
6: /cgn2_6/ptodata/1/ina/backfiles1.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	22.2	1.3	47	1	US-08-334-177-5	Sequence 5, Appli
	2	22.2	1.3	47	5	PCT-US95-13830-5	Sequence 5, Appli
c	3	21.8	1.2	44	4	US-08-507-426C-9	Sequence 9, Appli
	4	21.4	1.2	46	1	US-08-171-389-156	Sequence 156, App
	5	21.4	1.2	46	1	US-08-123-936-156	Sequence 156, App
	6	21.4	1.2	46	2	US-08-475-228A-156	Sequence 156, App
	7	21.4	1.2	46	3	US-08-482-080A-156	Sequence 156, App
	8	21.4	1.2	46	4	US-09-354-947-156	Sequence 156, App
	9	21.4	1.2	46	5	PCT-US93-12388-156	Sequence 156, App
	10	20.6	1.2	45	1	US-08-233-009-33	Sequence 33, Appl
c	11	20.6	1.2	45	1	US-08-285-936-38	Sequence 38, Appl
c	12	20.6	1.2	45	1	US-08-487-860-38	Sequence 38, Appl
	13	20.4	1.2	48	2	US-08-850-049-47	Sequence 47, Appl
	14	20.4	1.2	48	2	US-08-050-478-47	Sequence 47, Appl
	15	20.4	1.2	48	4	US-09-414-117-47	Sequence 47, Appl
	16	20.4	1.2	48	4	US-09-678-437-47	Sequence 47, Appl
c	17	20.2	1.2	36	4	US-09-344-700-21	Sequence 21, Appl
	18	20.2	1.2	40	1	US-08-236-918A-13	Sequence 13, Appl
	19	20.2	1.2	40	4	US-09-150-864A-13	Sequence 13, Appl
	20	20.2	1.2	49	2	US-08-752-844-33	Sequence 33, Appl
	21	20.2	1.2	49	2	US-08-591-196-33	Sequence 33, Appl
	22	20	1.1	36	1	US-08-030-731A-18	Sequence 18, Appl
c	23	20	1.1	44	1	US-08-030-731A-19	Sequence 19, Appl
	24	19.6	1.1	39	3	US-09-100-664A-7	Sequence 7, Appli
	25	19.6	1.1	39	4	US-09-335-983-7	Sequence 7, Appli
c	26	19.6	1.1	42	4	US-09-358-972-235	Sequence 235, App
c	27	19.6	1.1	42	4	US-09-430-615-25	Sequence 25, Appl
c	28	19.6	1.1	45	1	US-08-078-683A-27	Sequence 27, Appl
	29	19.6	1.1	45	1	US-08-206-176-19	Sequence 19, Appl
	30	19.6	1.1	50	1	US-08-171-389-582	Sequence 582, App
	31	19.6	1.1	50	1	US-08-123-936-582	Sequence 582, App
	32	19.6	1.1	50	2	US-08-475-228A-582	Sequence 582, App
	33	19.6	1.1	50	3	US-08-482-080A-582	Sequence 582, App
	34	19.6	1.1	50	4	US-09-354-947-582	Sequence 582, App
	35	19.6	1.1	50	5	PCT-US93-12388-582	Sequence 582, App
	36	19.2	1.1	40	4	US-09-375-318-11	Sequence 11, Appl
	37	19.2	1.1	48	2	US-08-752-844-38	Sequence 38, Appl
	38	19.2	1.1	48	2	US-08-591-196-38	Sequence 38, Appl
	39	19.2	1.1	50	3	US-08-732-218-5	Sequence 5, Appli
	40	19	1.1	35	5	PCT-US93-00031-5	Sequence 5, Appli
	41	19	1.1	45	1	US-08-176-412-5	Sequence 5, Appli
c	42	19	1.1	45	1	US-08-176-412-6	Sequence 6, Appli
	43	19	1.1	45	2	US-08-683-262B-32	Sequence 32, Appl
	44	19	1.1	45	2	US-08-555-268A-5	Sequence 5, Appli
c	45	19	1.1	45	2	US-08-555-268A-6	Sequence 6, Appli

ALIGNMENTS

```
RESULT 1
US-08-334-177-5
; Sequence 5, Application US/08334177
; Patent No. 5696086
;  GENERAL INFORMATION:
;    APPLICANT:  Avraham, Hava Karsenty
;    APPLICANT:  Godowski, Paul J.
;    TITLE OF INVENTION:  Methods and Kits Using Macrophage Stimulating Protein
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  5.25 inch, 360 Kb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  patin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/334,177
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:
;      FILING DATE:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Marschang, Diane L.
;      REGISTRATION NUMBER:  35,600
;      REFERENCE/DOCKET NUMBER:  912
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-5416
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 bases
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-08-334-177-5

  Query Match             1.3%;  Score 22.2;  DB 1;  Length 47;
  Best Local Similarity   69.8%;  Pred. No. 5.8e+03;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy     241 GGCGGCAGTGACCCTGGAGAGGCCCCCACACGTGCTGCTCCTG 283
           ||  | |   ||| ||| | ||| ||||| | ||||||| |||
```

```
Db      2 GGACGAATCCACCATGGGGTGGCTCCCACTCCTGCTGCTTCTG 44


RESULT 2
PCT-US95-13830-5
; Sequence 5, Application PC/TUS9513830
;  GENERAL INFORMATION:
;    APPLICANT:  Genentech, Inc.
;    APPLICANT:  New England Deaconess Hospital Corp.
;    TITLE OF INVENTION:  Methods and Kits Using Macrophage Stimulating
;    TITLE OF INVENTION:  Protein
;    NUMBER OF SEQUENCES:  10
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genentech, Inc.
;      STREET:  460 Point San Bruno Blvd
;      CITY:  South San Francisco
;      STATE:  California
;      COUNTRY:  USA
;      ZIP:  94080
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  3.5 inch, 1.44 Mb floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WinPatin (Genentech)
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US95/13830
;      FILING DATE:
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Marschang, Diane L.
;      REGISTRATION NUMBER:  35,600
;      REFERENCE/DOCKET NUMBER:  P0912PCT
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  415/225-5416
;      TELEFAX:  415/952-9881
;      TELEX:  910/371-7168
;  INFORMATION FOR SEQ ID NO:  5:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  47 base pairs
;      TYPE:  Nucleic Acid
;      STRANDEDNESS:  Single
;      TOPOLOGY:  Linear
PCT-US95-13830-5

  Query Match             1.3%;  Score 22.2;  DB 5;  Length 47;
  Best Local Similarity   69.8%;  Pred. No. 5.8e+03;
  Matches   30;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy    241 GGCGGCAGTGACCCTGGAGAGGCCCCCACACGTGCTGCTCCTG 283
          ||  | |   ||| ||| | ||| ||||| | ||||||| |||
Db      2 GGACGAATCCACCATGGGGTGGCTCCCACTCCTGCTGCTTCTG 44


RESULT 3
US-08-507-426C-9/c
; Sequence 9, Application US/08507426C
; Patent No. 6265634
; GENERAL INFORMATION:
;  APPLICANT:  Lenee, Philippe
;  TITLE OF INVENTION:  POLYRIBOZYME CAPABLE OF CONFERRING ON PLANTS RESISTANCE
;  TITLE OF INVENTION:  TO VIRUSES AND RESISTANT PLANTS PRODUCING THIS
;  TITLE OF INVENTION:  POLYRIBOZYME
;  FILE REFERENCE:  43944-A-PCT-US
;  CURRENT APPLICATION NUMBER: US/08/507,426C
;  CURRENT FILING DATE:  1995-10-25
;  PRIOR APPLICATION NUMBER:  43944-A-PCT-US
;  PRIOR FILING DATE:  1995-10-25
;  NUMBER OF SEQ ID NOS:  14
;  SOFTWARE:  PatentIn Ver. 2.1
; SEQ ID NO 9
;   LENGTH: 44
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:Synthetic
;   OTHER INFORMATION: ribozymes and portions thereof
US-08-507-426C-9

  Query Match             1.2%;  Score 21.8;  DB 4;  Length 44;
  Best Local Similarity   70.7%;  Pred. No. 7.2e+03;
  Matches   29;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy   1559 CGTCGATGCCTGACTCAGGCAGGCCAGCTTTCCGCGTGGTG 1599
          | | | ||   ||||||||||  || ||||| ||| || ||
Db     42 CCTTGGTGGACGACTCAGGCACTCCTGCTTTGCGCCTGCTG 2


RESULT 4
US-08-171-389-156
; Sequence 156, Application US/08171389
; Patent No. 5578444
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/171,389
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175/G19P3
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human H1 histone gene FNC16
US-08-171-389-156
```

```
  Query Match             1.2%;  Score 21.4;  DB 1;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 9.3e+03;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 5
US-08-123-936-156
; Sequence 156, Application US/08123936
; Patent No. 5726014
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    TITLE OF INVENTION:  Screening Assay for the Detection of
;    TITLE OF INVENTION:  DNA-Binding Molecules
;    NUMBER OF SEQUENCES:  640
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/123,936
;      FILING DATE:
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0075.32/G19P2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human H1 histone gene FNC16
US-08-123-936-156

  Query Match             1.2%;  Score 21.4;  DB 1;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 9.3e+03;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 6
US-08-475-228A-156
; Sequence 156, Application US/08475228A
; Patent No. 5869241
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/475,228A
;      FILING DATE:  06-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Stratford, Carol A.
;      REGISTRATION NUMBER:  34,444
;      REFERENCE/DOCKET NUMBER:  4600-0175.21/G19P3D2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human H1 histone gene FNC16
US-08-475-228A-156

  Query Match             1.2%;  Score 21.4;  DB 2;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 9.3e+03;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 7
US-08-482-080A-156
; Sequence 156, Application US/08482080A
; Patent No. 6010849
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
```

```
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/482,080A
;      FILING DATE:  07-JUN-1995
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human H1 histone gene FNC16
US-08-482-080A-156

  Query Match             1.2%;  Score 21.4;  DB 3;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 9.3e+03;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  |||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 8
US-09-354-947-156
; Sequence 156, Application US/09354947
; Patent No. 6384208
;  GENERAL INFORMATION:
;    APPLICANT:  Edwards, Cynthia A.
;    APPLICANT:  Cantor, Charles R.
;    APPLICANT:  Andrews, Beth M.
;    APPLICANT:  Turin, Lisa M.
;    APPLICANT:  Fry, Kirk E.
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  664
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/354,947
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/482,080
;      FILING DATE:  07-JUN-1995
;      APPLICATION NUMBER:  US 08/171,389
;      FILING DATE:  20-DEC-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/723,618
;      FILING DATE:  27-JUN-1991
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/081,070
;      FILING DATE:  22-JUN-1993
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Brady, John F.
;      REGISTRATION NUMBER:  39,118
;      REFERENCE/DOCKET NUMBER:  4600-0175.20/G19P3D1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (650) 324-0880
;      TELEFAX:  (650) 324-0960
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human H1 histone gene FNC16
US-09-354-947-156

  Query Match             1.2%;  Score 21.4;  DB 4;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 9.3e+03;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  |||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 9
PCT-US93-12388-156
; Sequence 156, Application PC/TUS9312388
;  GENERAL INFORMATION:
;    APPLICANT:
;    TITLE OF INVENTION:  Sequence-Directed DNA Binding
;    TITLE OF INVENTION:  Molecules, Compositions and Methods
;    NUMBER OF SEQUENCES:  641
```

```
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Genelabs Technologies, Inc.
;      STREET:  505 Penobscot Drive
;      CITY:  Redwood City
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  94063
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/12388
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/123,936
;      FILING DATE:  17-SEP-1993
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/996,783
;      FILING DATE:  23-DEC-1992
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Fabian, Gary R.
;      REGISTRATION NUMBER:  33,875
;      REFERENCE/DOCKET NUMBER:  4600-0175.41/G19PCT2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (415) 324-0880
;      TELEFAX:  (415) 324-0960
;  INFORMATION FOR SEQ ID NO:  156:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  46 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  double
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
;    HYPOTHETICAL:  NO
;    ORIGINAL SOURCE:
;      INDIVIDUAL ISOLATE:  Human H1 histone gene FNC16
PCT-US93-12388-156

  Query Match              1.2%;  Score 21.4;  DB 5;  Length 46;
  Best Local Similarity   71.8%;  Pred. No. 9.3e+03;
  Matches   28;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1641 GCGGCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCC 1679
          ||||  |   ||| ||  ||||  ||||||||||| ||
Db      2 GCGGTGGATTGGACGCTCCACCAATCACAGGGCAGCGCC 40


RESULT 10
US-08-233-009-33
; Sequence 33, Application US/08233009
; Patent No. 5646156
;  GENERAL INFORMATION:
;    APPLICANT:  Jacobson, Marlene A
;    APPLICANT:  Johnson, Robert G
;    APPLICANT:  Salvatore, Christopher A
;    TITLE OF INVENTION:  INHIBITION OF EOSINOPHIL
;    TITLE OF INVENTION:  ACTIVATION THROUGH A3 ADENOSINE RECEPTOR ANTAGONISM
;    NUMBER OF SEQUENCES:  56
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merck & Co., Inc.
;      STREET:  P.O.Box 2000
;      CITY:  Rahway
;      STATE:  New Jersey
;      COUNTRY:  United States
;      ZIP:  07065
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/233,009
;      FILING DATE:  25-APR-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Bencen, Gerard H
;      REGISTRATION NUMBER:  35,746
;      REFERENCE/DOCKET NUMBER:  19219
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (908) 594-3901
;      TELEFAX:  (908)594-4720
;  INFORMATION FOR SEQ ID NO:  33:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  both
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  cDNA
;    HYPOTHETICAL:  NO
;    ANTI-SENSE:  YES
US-08-233-009-33

  Query Match              1.2%;  Score 20.6;  DB 1;  Length 45;
  Best Local Similarity   67.4%;  Pred. No. 1.5e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    737 CCTGCACCGCCATCCGGGAAGTGTCCCTGCTCAAGGACCTCAA 779
          |||||||| ||| | |   ||   |||||  ||||| | || |
Db      2 CCTGCACCACCAACTGCTTAGCCCCCCTGGCCAAGGTCATCCA 44


RESULT 11
US-08-285-936-38/c
; Sequence 38, Application US/08285936
; Patent No. 5728821
;  GENERAL INFORMATION:
;    APPLICANT:  Yelton, Dale
;    APPLICANT:  Glaser, Scott
;    APPLICANT:  Huse, William
;    APPLICANT:  Rosok, Mae J.
;    TITLE OF INVENTION:  No. 5728821el Mutant BR96 Antibodies and
;    TITLE OF INVENTION:  Functional Equivalents Reactive With Human Carcinomas
;    NUMBER OF SEQUENCES:  57
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  11150 Santa Monica Bldv., Suite 400
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90025-3395
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/285,936
;      FILING DATE:  04-AUG-1994
;      CLASSIFICATION:  424
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adriano, Sarah B.
;      REGISTRATION NUMBER:  34,470
;      REFERENCE/DOCKET NUMBER:  30436.16US01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  310-445-1140
;      TELEFAX:  310-445-9031
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
```

```
US-08-285-936-38

  Query Match             1.2%;  Score 20.6;  DB 1;  Length 45;
  Best Local Similarity   67.4%;  Pred. No. 1.5e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    328 ATTGTGCACGAGGACTTGAAGATGGGGTCTGATGGGGAGAGTG 370
          | |||||| |||| || || || |||| ||| |  |   | ||
Db     44 ACTGTGCAAGAGGCCTGGACGACGGGGCCTGGTTTGCTTACTG 2


RESULT 12
US-08-487-860-38/c
; Sequence 38, Application US/08487860
; Patent No. 5792456
;  GENERAL INFORMATION:
;    APPLICANT:  Yelton, Dale
;    APPLICANT:  Glaser, Scott
;    APPLICANT:  Huse, William
;    APPLICANT:  Rosok, Mae J.
;    TITLE OF INVENTION:  No. 5792456el Mutant BR96 Antibodies and
;    TITLE OF INVENTION:  Functional Equivalents Reactive With Human Carcinomas
;    NUMBER OF SEQUENCES:  62
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Merchant & Gould
;      STREET:  11150 Santa Monica Bldv., Suite 400
;      CITY:  Los Angeles
;      STATE:  CA
;      COUNTRY:  USA
;      ZIP:  90025-3395
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/487,860
;      FILING DATE:  07-JUN-1995
;      CLASSIFICATION:  530
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  Adriano, Sarah B.
;      REGISTRATION NUMBER:  34,470
;      REFERENCE/DOCKET NUMBER:  30436.16US01
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  310-445-1140
;      TELEFAX:  310-445-9031
;  INFORMATION FOR SEQ ID NO:  38:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  45 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
;    MOLECULE TYPE:  DNA (genomic)
US-08-487-860-38

  Query Match             1.2%;  Score 20.6;  DB 1;  Length 45;
  Best Local Similarity   67.4%;  Pred. No. 1.5e+04;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    328 ATTGTGCACGAGGACTTGAAGATGGGGTCTGATGGGGAGAGTG 370
          | |||||| |||| || || || |||| ||| |  |   | ||
Db     44 ACTGTGCAAGAGGCCTGGACGACGGGGCCTGGTTTGCTTACTG 2


RESULT 13
US-08-850-049-47
; Sequence 47, Application US/08850049
; Patent No. 5965726
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/850,049
;      FILING DATE:  02-MAY-1997
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-08-850-049-47

  Query Match             1.2%;  Score 20.4;  DB 2;  Length 48;
  Best Local Similarity   71.1%;  Pred. No. 1.7e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    325 GAGATTGTGCACGAGGACTTGAAGATGGGGTCTGATGG 362
          ||||  |||| || | | |||||| ||||  |||||
Db      7 GAGACGGTGCCCGTGAAGTTGAAGCCGGGGATGGATGG 44


RESULT 14
US-08-050-478-47
; Sequence 47, Application US/08050478
; Patent No. 5972596
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
```

```
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:  435
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:  435
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-08-050-478-47

  Query Match             1.2%;  Score 20.4;  DB 2;  Length 48;
  Best Local Similarity   71.1%;  Pred. No. 1.7e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    325 GAGATTGTGCACGAGGACTTGAAGATGGGGTCTGATGG 362
          ||||  |||| || | | ||||||  ||||   |||||
Db      7 GAGACGGTGCCCGTGAAGTTGAAGCCGGGGATGGATGG 44


RESULT 15
US-09-414-117-47
; Sequence 47, Application US/09414117
; Patent No. 6291664
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/414,117
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/850,049
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
;      TOPOLOGY:  LINEAR
US-09-414-117-47

  Query Match             1.2%;  Score 20.4;  DB 4;  Length 48;
  Best Local Similarity   71.1%;  Pred. No. 1.7e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    325 GAGATTGTGCACGAGGACTTGAAGATGGGGTCTGATGG 362
          ||||  |||| || | | ||||||  ||||   |||||
Db      7 GAGACGGTGCCCGTGAAGTTGAAGCCGGGGATGGATGG 44


Search completed: March  4, 2003, 00:52:00
Job time : 80 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 22:39:05 ; Search time 133 Seconds
(without alignments)
8182.066 Million cell updates/sec

Title: US-10-017-621-3
Perfect score: 1745
Sequence: 1 tggaagcagcgtaaaggatg.....gttcacctgcccacttgtcc 1745

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 460893 seqs, 311809382 residues

Total number of hits satisfying chosen parameters: 254638

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : Published_Applications_NA:*
1: /cgn2_6/ptodata/2/pubpna/US07_PUBCOMB.seq:*
2: /cgn2_6/ptodata/2/pubpna/PCT_NEW_PUB.seq:*
3: /cgn2_6/ptodata/2/pubpna/US06_NEW_PUB.seq:*
4: /cgn2_6/ptodata/2/pubpna/US06_PUBCOMB.seq:*
5: /cgn2_6/ptodata/2/pubpna/US07_NEW_PUB.seq:*
6: /cgn2_6/ptodata/2/pubpna/PCTUS_PUBCOMB.seq:*
7: /cgn2_6/ptodata/2/pubpna/US08_NEW_PUB.seq:*
8: /cgn2_6/ptodata/2/pubpna/US08_PUBCOMB.seq:*
9: /cgn2_6/ptodata/2/pubpna/US09_NEW_PUB.seq:*
10: /cgn2_6/ptodata/2/pubpna/US09_PUBCOMB.seq:*
11: /cgn2_6/ptodata/2/pubpna/US10_NEW_PUB.seq:*
12: /cgn2_6/ptodata/2/pubpna/US10_PUBCOMB.seq:*
13: /cgn2_6/ptodata/2/pubpna/US60_NEW_PUB.seq:*
14: /cgn2_6/ptodata/2/pubpna/US60_PUBCOMB.seq:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	21.8	1.2	48	9	US-10-054-444-6	Sequence 6, Appli
	2	21.6	1.2	31	10	US-09-801-274-752	Sequence 752, App
	3	21.4	1.2	45	9	US-10-029-413A-25	Sequence 25, Appl
	4	21	1.2	31	10	US-09-801-274-94	Sequence 94, Appl
	5	20.4	1.2	48	9	US-09-943-722-47	Sequence 47, Appl
	6	20	1.1	45	10	US-09-147-142-11	Sequence 11, Appl
c	7	20	1.1	45	10	US-09-147-142-12	Sequence 12, Appl
	8	19.8	1.1	40	10	US-09-263-959-758	Sequence 758, App
c	9	19.6	1.1	42	10	US-09-790-417-235	Sequence 235, App
c	10	19.4	1.1	45	12	US-10-073-256-27	Sequence 27, Appl
c	11	19.4	1.1	45	12	US-10-073-256-35	Sequence 35, Appl
	12	19.2	1.1	31	10	US-09-801-274-517	Sequence 517, App
c	13	19.2	1.1	46	10	US-09-263-959-121	Sequence 121, App
c	14	19	1.1	43	9	US-09-376-940-23	Sequence 23, Appl
	15	19	1.1	45	10	US-09-818-066-32	Sequence 32, Appl
c	16	19	1.1	47	9	US-10-118-231-9	Sequence 9, Appli
	17	18.8	1.1	48	9	US-09-840-277-104	Sequence 104, App
	18	18.8	1.1	48	9	US-09-840-277-105	Sequence 105, App
c	19	18.8	1.1	48	10	US-09-753-436-67	Sequence 67, Appl
c	20	18.8	1.1	49	9	US-09-840-277-107	Sequence 107, App
	21	18.6	1.1	25	10	US-09-866-108-15295	Sequence 15295, A
	22	18.6	1.1	48	9	US-09-864-785-3433	Sequence 3433, Ap
	23	18.4	1.1	41	9	US-09-938-842A-5256	Sequence 5256, Ap
	24	18.4	1.1	48	10	US-09-761-534A-18	Sequence 18, Appl
	25	18.2	1.0	36	9	US-10-219-248-31	Sequence 31, Appl
	26	18.2	1.0	36	9	US-10-219-247-31	Sequence 31, Appl
	27	18.2	1.0	36	10	US-09-855-722-31	Sequence 31, Appl
	28	18.2	1.0	50	9	US-09-943-722-9	Sequence 9, Appli
c	29	18	1.0	42	10	US-09-790-417-233	Sequence 233, App
	30	18	1.0	43	12	US-10-043-142-4	Sequence 4, Appli
	31	18	1.0	46	9	US-10-026-914-8	Sequence 8, Appli
	32	18	1.0	46	9	US-10-026-914-10	Sequence 10, Appl
	33	18	1.0	46	9	US-10-026-914-14	Sequence 14, Appl
	34	18	1.0	46	9	US-10-026-914-16	Sequence 16, Appl
	35	18	1.0	48	9	US-09-864-785-3226	Sequence 3226, Ap
c	36	17.8	1.0	35	9	US-10-051-989-2	Sequence 2, Appli
c	37	17.8	1.0	35	9	US-09-861-097-2	Sequence 2, Appli
	38	17.8	1.0	42	10	US-09-865-807-25	Sequence 25, Appl
c	39	17.8	1.0	45	12	US-10-073-256-24	Sequence 24, Appl
c	40	17.8	1.0	45	12	US-10-073-256-25	Sequence 25, Appl
c	41	17.8	1.0	45	12	US-10-073-256-33	Sequence 33, Appl
c	42	17.8	1.0	45	12	US-10-073-256-34	Sequence 34, Appl
	43	17.8	1.0	48	9	US-09-864-785-3267	Sequence 3267, Ap
	44	17.8	1.0	48	9	US-09-864-785-3311	Sequence 3311, Ap
	45	17.8	1.0	48	9	US-09-864-785-3491	Sequence 3491, Ap

ALIGNMENTS

RESULT 1
US-10-054-444-6
; Sequence 6, Application US/10054444
; Patent No. US20020164342A1
; GENERAL INFORMATION:
; APPLICANT: Guyre, Paul M.
; APPLICANT: Goldstein, Joel
; APPLICANT: Wu, Zining
; APPLICANT: Sun, Wanwen
; TITLE OF INVENTION: Recombinant Cat Allergen, Fel dI, Expressed in
; TITLE OF INVENTION: Baculovirus for Diagnosis and Treatment of Cat Allergy
; FILE REFERENCE: DC-0118
; CURRENT APPLICATION NUMBER: US/10/054,444
; CURRENT FILING DATE: 2002-01-22
; PRIOR APPLICATION NUMBER: EARLIER APPLICATION NUMBER: 09/410,963
; PRIOR FILING DATE: EARLIER FILING DATE: 1999-10-05
; NUMBER OF SEQ ID NOS: 6
; SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 6
; LENGTH: 48
; TYPE: DNA
; ORGANISM: Artificial Sequence
; FEATURE:
; OTHER INFORMATION: Description of Artificial Sequence:Synthetic
US-10-054-444-6

Query Match 1.2%; Score 21.8; DB 9; Length 48;
Best Local Similarity 70.7%; Pred. No. 9.5e+03;
Matches 29; Conservative 0; Mismatches 12; Indels 0; Gaps 0;

```
Qy     962 AGAAGGTGCTACACCGAGACCTCAAGCCCCAGAACCTGCTC 1002
           ||||  | || ||||   |||||   | |||||||| |||
Db       2 AGAACCTCCTCCACCAGAACCTCCTCCACCAGAACCTCCTC 42
```

RESULT 2
US-09-801-274-752
; Sequence 752, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
; APPLICANT: Cargill, Michele
; APPLICANT: Ireland, James S.

```
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 752
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-752

  Query Match             1.2%;  Score 21.6;  DB 10;  Length 31;
  Best Local Similarity   80.0%;  Pred. No. 8.8e+03;
  Matches   24;  Conservative    1;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    979 GACCTCAAGCCCCAGAACCTGCTCATCAAC 1008
          ||| |||||||||||:||||||||  |  ||
Db      2 GACATCAAGCCCCAKAACCTGCTGGTGGAC 31


RESULT 3
US-10-029-413A-25
; Sequence 25, Application US/10029413A
; Patent No. US20020165353A1
; GENERAL INFORMATION:
;  APPLICANT: Malouf, Nadia
;  APPLICANT:  Nichols, Timothy C.
;  TITLE OF INVENTION: Purified and Isolated Platelet Calcium Channel Nucleic Acids and
;  TITLE OF INVENTION:  Polypeptides and Therapeutic and Screening Methods Using Same
;  FILE REFERENCE: 421/29
;  CURRENT APPLICATION NUMBER: US/10/029,413A
;  CURRENT FILING DATE:  2002-03-21
;  NUMBER OF SEQ ID NOS: 29
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 25
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Homo sapiens
;   FEATURE:
;   NAME/KEY: partial_cDNA
;   LOCATION: (1)..(45)
;   OTHER INFORMATION:
US-10-029-413A-25

  Query Match             1.2%;  Score 21.4;  DB 9;  Length 45;
  Best Local Similarity   80.6%;  Pred. No. 1.2e+04;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    907 AACGTGAAACTGTTCCTGTTCCAGCTGCTCC 937
          || |||||| ||| ||||| ||| |||||  |
Db      8 AAAGTGAAAATGTCCCTGTCCCAACTGCTAC 38


RESULT 4
US-09-801-274-94
; Sequence 94, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 94
;   LENGTH: 31
;   TYPE: DNA
;   ORGANISM: Homo sapiens
US-09-801-274-94

  Query Match             1.2%;  Score 21;  DB 10;  Length 31;
  Best Local Similarity   77.4%;  Pred. No. 1.3e+04;
  Matches   24;  Conservative    1;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    577 GTCAGCCTATCTGAGATTGGCTTTGGGAAAC 607
          | |  ||| || || :|||||||||||||||
Db      1 GCCTCCCTGTCAGACMTTGGCTTTGGGAAAC 31


RESULT 5
US-09-943-722-47
; Sequence 47, Application US/09943722
; Publication No. US20020192660A1
;  GENERAL INFORMATION:
;    APPLICANT:
;    APPLICANT:
;    APPLICANT:
;    TITLE OF INVENTION:  METHOD OF ELIMINATING
;    TITLE OF INVENTION:  INHIBITORY/INSTABILITY REGIONS OF mRNA
;    NUMBER OF SEQUENCES:  130
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  MORGAN & FINNEGAN
;      STREET:  345 PARK AVENUE
;      CITY:  NEW YORK
;      STATE:  NEW YORK
;      COUNTRY:  USA
;      ZIP:  10154
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  FLOPPY DISK
;      COMPUTER:  IBM PC COMPATIBLE
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  WORDPERFECT 5.1
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/943,722
;      FILING DATE:
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  08/850,049
;      FILING DATE:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 08/050,478
;      FILING DATE:  26-OCT-1994
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  PCT/US93/02908
;      FILING DATE:  29-MAR-1993
;      CLASSIFICATION:
;    PRIOR APPLICATION DATA:
;      APPLICATION NUMBER:  US 07/858,747
;      FILING DATE:  27-MAR-1992
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  MORRY, MARY J.
;      REGISTRATION NUMBER:  34,398
;      REFERENCE/DOCKET NUMBER:  2026-4006US1
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (212)758-4800
;      TELEFAX:  (212)751-6849
;  INFORMATION FOR SEQ ID NO:  47:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  48 BASE PAIRS
;      TYPE:  NUCLEIC ACID
;      STRANDEDNESS:  SINGLE
```

```
;      TOPOLOGY:  LINEAR
US-09-943-722-47

  Query Match             1.2%;  Score 20.4;  DB 9;  Length 48;
  Best Local Similarity   71.1%;  Pred. No. 2.3e+04;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy    325 GAGATTGTGCACGAGGACTTGAAGATGGGGTCTGATGG 362
          ||||  |||| || | | ||||||  ||||   |||||
Db      7 GAGACGGTGCCCGTGAAGTTGAAGCCGGGGATGGATGG 44


RESULT 6
US-09-147-142-11
; Sequence 11, Application US/09147142
; Patent No. US20020018749A1
; GENERAL INFORMATION:
;  APPLICANT: HUDSON, Peter John
;  APPLICANT: KORTT, Alex Andrew
;  APPLICANT: IRVING, Robert Alexander
;  APPLICANT: ATWELL, John Leslie
;  TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
;  FILE REFERENCE: 016786/0212
;  CURRENT APPLICATION NUMBER: US/09/147,142
;  CURRENT FILING DATE: 1999-03-05
;  EARLIER APPLICATION NUMBER: PCT/AU98/00212
;  EARLIER FILING DATE: 1998-03-26
;  EARLIER APPLICATION NUMBER: AU PO 5917
;  EARLIER FILING DATE: 1997-03-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 11
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide used to insert codon between VH
;   OTHER INFORMATION: and VL domains of NC10 scFv-0
US-09-147-142-11

  Query Match             1.1%;  Score 20;  DB 10;  Length 45;
  Best Local Similarity   65.9%;  Pred. No. 2.8e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    643 GGTACCTATGCCACCGTCTACAAAGGCAAAAGCAAGCTCACAGA 686
          || |||   | |||||||| |   ||  | | | |||||||| |
Db      1 GGGACCACGGTCACCGTCTCCGGTGGTGATATCGAGCTCACACA 44


RESULT 7
US-09-147-142-12/c
; Sequence 12, Application US/09147142
; Patent No. US20020018749A1
; GENERAL INFORMATION:
;  APPLICANT: HUDSON, Peter John
;  APPLICANT: KORTT, Alex Andrew
;  APPLICANT: IRVING, Robert Alexander
;  APPLICANT: ATWELL, John Leslie
;  TITLE OF INVENTION: HIGH AVIDITY POLYVALENT AND POLYSPECIFIC REAGENTS
;  FILE REFERENCE: 016786/0212
;  CURRENT APPLICATION NUMBER: US/09/147,142
;  CURRENT FILING DATE: 1999-03-05
;  EARLIER APPLICATION NUMBER: PCT/AU98/00212
;  EARLIER FILING DATE: 1998-03-26
;  EARLIER APPLICATION NUMBER: AU PO 5917
;  EARLIER FILING DATE: 1997-03-27
;  NUMBER OF SEQ ID NOS: 32
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 12
;   LENGTH: 45
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence: synthetic
;   OTHER INFORMATION: oligonucleotide used to insert codon between VH
;   OTHER INFORMATION: and VL domains of NC10 scFv-0
US-09-147-142-12

  Query Match             1.1%;  Score 20;  DB 10;  Length 45;
  Best Local Similarity   65.9%;  Pred. No. 2.8e+04;
  Matches   29;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy    643 GGTACCTATGCCACCGTCTACAAAGGCAAAAGCAAGCTCACAGA 686
          || |||   | |||||||| |   ||  | | | |||||||| |
Db     45 GGGACCACGGTCACCGTCTCCGGTGGTGATATCGAGCTCACACA 2


RESULT 8
US-09-263-959-758
; Sequence 758, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
;      OPERATING SYSTEM:  PC-DOS/MS-DOS
;      SOFTWARE:  PatentIn Release #1.0, Version #1.25
;    CURRENT APPLICATION DATA:
;      APPLICATION NUMBER:  US/09/263,959
;      FILING DATE:  05-MAR-1999
;      CLASSIFICATION:
;    ATTORNEY/AGENT INFORMATION:
;      NAME:  McMasters, David D.
;      REGISTRATION NUMBER:  33,963
;      REFERENCE/DOCKET NUMBER:  920010.426C2
;    TELECOMMUNICATION INFORMATION:
;      TELEPHONE:  (206) 622-4900
;      TELEFAX:  (206) 682-6031
;  INFORMATION FOR SEQ ID NO:  758:
;    SEQUENCE CHARACTERISTICS:
;      LENGTH:  40 base pairs
;      TYPE:  nucleic acid
;      STRANDEDNESS:  single
;      TOPOLOGY:  linear
US-09-263-959-758

  Query Match             1.1%;  Score 19.8;  DB 10;  Length 40;
  Best Local Similarity   77.4%;  Pred. No. 3e+04;
  Matches   24;  Conservative    0;  Mismatches    7;  Indels    0;  Gaps    0;

Qy   1689 CTTCCCTGCTTACTCTCTGCCTACCTGCCTG 1719
          ||||| |  || || ||| ||| ||||||||
Db      3 CTTCCTTCTTTCCTTTCTTCCTGCCTGCCTG 33


RESULT 9
US-09-790-417-235/c
; Sequence 235, Application US/09790417
; Patent No. US20010031470A1
; GENERAL INFORMATION:
;  APPLICANT: Shultz, John W
;  APPLICANT:  Lewis, Martin K.
;  APPLICANT:  Lieppe, Donna
```

```
;  APPLICANT:  Mandrekar, Michelle
;  APPLICANT:  Kephart, Daniel
;  APPLICANT:  Rhodes, Richard B.
;  APPLICANT:  Andrews, Christine A.
;  APPLICANT:  Hartnett, James R.
;  APPLICANT:  Gu, Trent
;  APPLICANT:  Olson, Ryan J.
;  APPLICANT:  Wood, Keith W.
;  APPLICANT:  Welch, Roy
;  TITLE OF INVENTION: Nucleic Acid Detection
;  FILE REFERENCE: Pro-103 6868/75528
;  CURRENT APPLICATION NUMBER: US/09/790,417
;  CURRENT FILING DATE:  2001-02-22
;  PRIOR APPLICATION NUMBER: 09/358,972
;  PRIOR FILING DATE: 1999-07-21
;  PRIOR APPLICATION NUMBER: 09/042,287
;  PRIOR FILING DATE: 1998-03-13
;  NUMBER OF SEQ ID NOS: 290
;  SOFTWARE: PatentIn Ver. 2.0
; SEQ ID NO 235
;    LENGTH: 42
;    TYPE: DNA
;    ORGANISM: Artificial Sequence
;    FEATURE:
;    OTHER INFORMATION: Description of Artificial Sequence:probe for oligo
;    OTHER INFORMATION: 54
US-09-790-417-235

  Query Match             1.1%;  Score 19.6;  DB 10;  Length 42;
  Best Local Similarity   66.7%;  Pred. No. 3.5e+04;
  Matches   28;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy    759 GTCCCTGCTCAAGGACCTCAAACACGCCAACATCGTTACGCT 800
          || |||| | || || |||| |||| | |||| | |||
Db     42 GTACCTGGTAAATGAACTCACCCACGATATCATCACCAAGCT 1


RESULT 10
US-10-073-256-27/c
; Sequence 27, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
;  APPLICANT: Kreiswirth, Barry N
;  APPLICANT:  Nadich, Steven M
;  TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
;  FILE REFERENCE: 19124.0002
;  CURRENT APPLICATION NUMBER: US/10/073,256
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 27
;    LENGTH: 45
;    TYPE: DNA
;    ORGANISM: Enterococcus faecalis
US-10-073-256-27

  Query Match             1.1%;  Score 19.4;  DB 12;  Length 45;
  Best Local Similarity   64.4%;  Pred. No. 4.1e+04;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1288 ATCCTGTCCAACGAGGAGTTCAAGACATACAACTACCCCAAGTAC 1332
          | || |   ||| || ||| | ||  | |  | |||| |||||||
Db     45 AACCAGCGAAACCAGTAGTACCAGCGAAAGTAGTACCTCAAGTAC 1


RESULT 11
US-10-073-256-35/c
; Sequence 35, Application US/10073256
; Patent No. US20020120408A1
; GENERAL INFORMATION:
;  APPLICANT: Kreiswirth, Barry N
;  APPLICANT:  Nadich, Steven M
;  TITLE OF INVENTION: System and Method for Tracking and Controlling Infections
;  FILE REFERENCE: 19124.0002
;  CURRENT APPLICATION NUMBER: US/10/073,256
;  CURRENT FILING DATE:  2002-02-13
;  NUMBER OF SEQ ID NOS: 80
;  SOFTWARE: PatentIn version 3.1
; SEQ ID NO 35
;    LENGTH: 45
;    TYPE: DNA
;    ORGANISM: Enterococcus faecalis
US-10-073-256-35

  Query Match             1.1%;  Score 19.4;  DB 12;  Length 45;
  Best Local Similarity   64.4%;  Pred. No. 4.1e+04;
  Matches   29;  Conservative    0;  Mismatches   16;  Indels    0;  Gaps    0;

Qy   1288 ATCCTGTCCAACGAGGAGTTCAAGACATACAACTACCCCAAGTAC 1332
          | || ||  ||| || | | |||| | |  |  ||| |||||||
Db     45 AACCAGTGAAACCAGTAATACAAGCGAAAGCAGCACCTCAAGTAC 1


RESULT 12
US-09-801-274-517
; Sequence 517, Application US/09801274
; Patent No. US20020032319A1
; GENERAL INFORMATION:
;  APPLICANT: Cargill, Michele
;  APPLICANT:  Ireland, James S.
;  APPLICANT:  Lander, Eric S.
;  TITLE OF INVENTION: HUMAN SINGLE NUCLEOTIDE POLYMORPHISMS
;  FILE REFERENCE: 2825.2009-001
;  CURRENT APPLICATION NUMBER: US/09/801,274
;  CURRENT FILING DATE:  2001-03-07
;  PRIOR APPLICATION NUMBER: US 60/187,510
;  PRIOR FILING DATE: 2000-03-07
;  PRIOR APPLICATION NUMBER: US 60/206,129
;  PRIOR FILING DATE: 2000-05-22
;  NUMBER OF SEQ ID NOS: 1802
;  SOFTWARE: FastSEQ for Windows Version 4.0
; SEQ ID NO 517
;    LENGTH: 31
;    TYPE: DNA
;    ORGANISM: Homo sapiens
US-09-801-274-517

  Query Match             1.1%;  Score 19.2;  DB 10;  Length 31;
  Best Local Similarity   80.8%;  Pred. No. 3.9e+04;
  Matches   21;  Conservative    1;  Mismatches    4;  Indels    0;  Gaps    0;

Qy   1190 CCACAGGCCGTCCCCTCTTTCCGGGC 1215
          ||||||||| |:|||||||   ||||
Db      5 CCACAGGCCTTYCCCTCTTCGTGGGC 30


RESULT 13
US-09-263-959-121/c
; Sequence 121, Application US/09263959
; Patent No. US20020150891A1
;  GENERAL INFORMATION:
;    APPLICANT:  Hood, Leroy E.
;    APPLICANT:  Rowen, Lee
;    APPLICANT:  Koop, Ben F.
;    TITLE OF INVENTION:  DIAGNOSTIC AND THERAPEUTIC COMPOSITIONS AND METHODS WHICH U
;    NUMBER OF SEQUENCES:  1279
;    CORRESPONDENCE ADDRESS:
;      ADDRESSEE:  Seed and Berry LLP
;      STREET:  6300 Columbia Center, 701 Fifth Avenue
;      CITY:  Seattle
;      STATE:  Washington
;      COUNTRY:  US
;      ZIP:  98104-7092
;    COMPUTER READABLE FORM:
;      MEDIUM TYPE:  Floppy disk
;      COMPUTER:  IBM PC compatible
```

```
;       OPERATING SYSTEM:  PC-DOS/MS-DOS
;       SOFTWARE:  PatentIn Release #1.0, Version #1.25
;     CURRENT APPLICATION DATA:
;       APPLICATION NUMBER:  US/09/263,959
;       FILING DATE:  05-MAR-1999
;       CLASSIFICATION:
;     ATTORNEY/AGENT INFORMATION:
;       NAME:  McMasters, David D.
;       REGISTRATION NUMBER:  33,963
;       REFERENCE/DOCKET NUMBER:  920010.426C2
;     TELECOMMUNICATION INFORMATION:
;       TELEPHONE:  (206) 622-4900
;       TELEFAX:  (206) 682-6031
;   INFORMATION FOR SEQ ID NO:  121:
;     SEQUENCE CHARACTERISTICS:
;       LENGTH:  46 base pairs
;       TYPE:  nucleic acid
;       STRANDEDNESS:  single
;       TOPOLOGY:  linear
US-09-263-959-121

  Query Match             1.1%;  Score 19.2;  DB 10;  Length 46;
  Best Local Similarity   75.0%;  Pred. No. 4.7e+04;
  Matches   24;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    319 GCACCAGAGATTGTGCACGAGGACTTGAAGAT 350
          ||||||||| || |||||  |||   | ||||
Db     46 GCACCAGAGTTTCTGCACAGGGAGGGGGAGAT 15


RESULT 14
US-09-376-940-23/c
; Sequence 23, Application US/09376940A
; Publication No. US20020192813A1
; GENERAL INFORMATION:
;  APPLICANT: Conner, Timothy W
;  APPLICANT: Santino, Colleen G
;  TITLE OF INVENTION: No. US20020192813A1el Plant Expression Vectors
;  FILE REFERENCE: monocot elements
;  CURRENT APPLICATION NUMBER: US/09/376,940A
;  CURRENT FILING DATE: 1999-08-18
;  EARLIER APPLICATION NUMBER: 60/097150
;  EARLIER FILING DATE: 1998-08-19
;  NUMBER OF SEQ ID NOS: 63
;  SOFTWARE: PatentIn Ver. 2.1
; SEQ ID NO 23
;   LENGTH: 43
;   TYPE: DNA
;   ORGANISM: Artificial Sequence
;   FEATURE:
;   OTHER INFORMATION: Description of Artificial Sequence:synthetic
US-09-376-940-23

  Query Match             1.1%;  Score 19;  DB 9;  Length 43;
  Best Local Similarity   65.1%;  Pred. No. 5.1e+04;
  Matches   28;  Conservative    0;  Mismatches   15;  Indels    0;  Gaps    0;

Qy     79 GGGCCCCGCGGCTCTGAGGTTGCTCGCGCGCCCCCGCCGATCG 121
          | | ||  |  | |   | |  ||||| |||| ||||||||||
Db     43 GTGTCCACCCACCCCTCGATCTCTCGCTCGCCGCCGCCGATCG 1


RESULT 15
US-09-818-066-32
; Sequence 32, Application US/09818066
; Patent No. US20020032307A1
;   GENERAL INFORMATION:
;       APPLICANT: Shuping Tong et al.
;       TITLE OF INVENTION: HEPADNAVIRUS RECEPTOR
;       NUMBER OF SEQUENCES: 75
;       CORRESPONDENCE ADDRESS:
;            ADDRESSEE: Fish & Richardson P.C.
;            STREET: 225 Franklin Street
;            CITY: Boston
;            STATE: Massachusetts
;            COUNTRY: U.S.A.
;            ZIP: 02110-2804
;       COMPUTER READABLE FORM:
;            MEDIUM TYPE: Floppy disk
;            COMPUTER: IBM PC compatible
;            OPERATING SYSTEM: PC-DOS/MS-DOS
;            SOFTWARE: PatentIn Release #1.0, Version #1.30
;       CURRENT APPLICATION DATA:
;            APPLICATION NUMBER: US/09/818,066
;            FILING DATE: 27-Mar-2001
;            CLASSIFICATION: <Unknown>
;       PRIOR APPLICATION DATA:
;            APPLICATION NUMBER: US 08/683,262
;            FILING DATE: 18-JUL-1996
;       ATTORNEY/AGENT INFORMATION:
;            NAME: Fraser, Janis K.
;            REGISTRATION NUMBER: 31,819
;            REFERENCE/DOCKET NUMBER: 00786/287002
;       TELECOMMUNICATION INFORMATION:
;            TELEPHONE: (617) 542-5070
;            TELEFAX: (617) 542-8906
;            TELEX: 200154
;   INFORMATION FOR SEQ ID NO: 32:
;       SEQUENCE CHARACTERISTICS:
;            LENGTH: 45 base pairs
;            TYPE: nucleic acid
;            STRANDEDNESS: single
;            TOPOLOGY: linear
;       SEQUENCE DESCRIPTION: SEQ ID NO: 32:
US-09-818-066-32

  Query Match             1.1%;  Score 19;  DB 10;  Length 45;
  Best Local Similarity   71.4%;  Pred. No. 5.2e+04;
  Matches   25;  Conservative    0;  Mismatches   10;  Indels    0;  Gaps    0;

Qy   1343 CCCTTTTGAGCCACGCACCCCGACTTGATAGCGAC 1377
          |||||||||||||||  |  ||    ||| |  |||
Db      2 CCCTTTTGAGCCACGAATTCCAGGATGAAACAGAC 36


Search completed: March  4, 2003, 00:54:25
Job time : 134 secs
```

GenCore version 5.1.4_p5_4578

OM nucleic - nucleic search, using sw model

Run on: March 3, 2003, 22:34:05 ; Search time 2628 Seconds
(without alignments)
10753.853 Million cell updates/sec

Title: US-10-017-621-3
Perfect score: 1745
Sequence: 1 tggaagcagcgtaaaggatg.....gttcacctgcccacttgtcc 1745

Scoring table: IDENTITY_NUC
Gapop 10.0 , Gapext 1.0

Searched: 16154066 seqs, 8097743376 residues

Total number of hits satisfying chosen parameters: 102860

Minimum DB seq length: 0
Maximum DB seq length: 50

Post-processing: Minimum Match 0%
Maximum Match 100%
Listing first 45 summaries

Database : EST:*
1: em_estba:*
2: em_esthum:*
3: em_estin:*
4: em_estmu:*
5: em_estov:*
6: em_estpl:*
7: em_estro:*
8: em_htc:*
9: gb_est1:*
10: gb_est2:*
11: gb_htc:*
12: gb_est3:*
13: gb_est4:*
14: gb_est5:*
15: em_estfun:*
16: em_estom:*
17: gb_gss:*
18: em_gss_hum:*
19: em_gss_inv:*
20: em_gss_pln:*
21: em_gss_vrt:*
22: em_gss_fun:*
23: em_gss_mam:*
24: em_gss_mus:*
25: em_gss_other:*
26: em_gss_pro:*
27: em_gss_rod:*

Pred. No. is the number of results predicted by chance to have a score greater than or equal to the score of the result being printed, and is derived by analysis of the total score distribution.

SUMMARIES

	Result No.	Score	% Query Match	Length	DB	ID	Description
	1	45	2.6	46	14	N78054	N78054 yv71g05.r1
c	2	28	1.6	28	14	R38968	R38968 yd07c08.s1
	3	23.2	1.3	48	10	AW247978	AW247978 2819657.5
	4	21.4	1.2	50	9	AU107934	AU107934 AU107934
	5	21.2	1.2	36	17	AZ346286	AZ346286 1M0081C01
	6	21	1.2	50	9	AU102877	AU102877 AU102877
c	7	21	1.2	50	9	AU105237	AU105237 AU105237
c	8	21	1.2	50	13	BM397711	BM397711 5009-0-35
	9	20.8	1.2	46	17	AZ993993	AZ993993 2M0279E13
	10	20.6	1.2	48	17	AZ311362	AZ311362 1M0026F16
c	11	20.6	1.2	50	9	AU106960	AU106960 AU106960
c	12	20.4	1.2	47	17	AZ331536	AZ331536 1M0059H04
	13	20.2	1.2	45	17	AZ985975	AZ985975 2M0268F01
c	14	20.2	1.2	49	9	AA204601	AA204601 mu25c05.r
c	15	20	1.1	40	9	AI475974	AI475974 tl96b06.x
	16	20	1.1	50	17	BH811451	BH811451 SALK_0586
	17	19.8	1.1	49	14	W39000	W39000 zb29b05.r1
c	18	19.4	1.1	50	9	AU104829	AU104829 AU104829
	19	19.2	1.1	44	12	BG422154	BG422154 602448881
	20	19.2	1.1	49	17	AZ450961	AZ450961 1M0250B05
c	21	19.2	1.1	50	9	AU103357	AU103357 AU103357
c	22	19.2	1.1	50	9	AU103358	AU103358 AU103358
c	23	19.2	1.1	50	9	AU103359	AU103359 AU103359
c	24	19.2	1.1	50	9	AU103361	AU103361 AU103361
	25	19.2	1.1	50	9	AU103381	AU103381 AU103381
	26	19.2	1.1	50	9	AU103915	AU103915 AU103915
	27	19.2	1.1	50	9	AU106349	AU106349 AU106349
c	28	19.2	1.1	50	14	T74703	T74703 yc60g05.s1
	29	19	1.1	43	9	AI591257	AI591257 tt75c06.x
	30	19	1.1	49	17	AZ423762	AZ423762 1M0203E22
	31	19	1.1	50	9	AU107320	AU107320 AU107320
	32	18.8	1.1	34	9	AA972479	AA972479 op42c10.s
c	33	18.8	1.1	43	17	BH857724	BH857724 SALK_0159
	34	18.8	1.1	46	17	BH790015	BH790015 SALK_0529
	35	18.8	1.1	50	9	AU102939	AU102939 AU102939
	36	18.8	1.1	50	9	AU103583	AU103583 AU103583
c	37	18.8	1.1	50	9	AU104587	AU104587 AU104587
	38	18.8	1.1	50	9	AU105918	AU105918 AU105918
c	39	18.6	1.1	44	17	TA165G05Q	AL473116 T. brucei
c	40	18.6	1.1	45	9	AI250043	AI250043 qx48f02.x
	41	18.6	1.1	49	17	AZ966392	AZ966392 2M0236B20
	42	18.6	1.1	50	9	AU103382	AU103382 AU103382
	43	18.6	1.1	50	9	AU103553	AU103553 AU103553
c	44	18.6	1.1	50	9	AU104162	AU104162 AU104162
c	45	18.6	1.1	50	13	BI910989	BI910989 603069394

ALIGNMENTS

RESULT 1
N78054
LOCUS N78054 46 bp mRNA linear EST 28-JAN-1997
DEFINITION yv71g05.r1 Soares fetal liver spleen 1NFLS Homo sapiens cDNA clone IMAGE:248216 5' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE PCTAIRE-1 (HUMAN);, mRNA sequence.
ACCESSION N78054
VERSION N78054.1 GI:1240755
KEYWORDS EST.
SOURCE human.
ORGANISM Homo sapiens
Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi; Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE 1 (bases 1 to 46)
AUTHORS Hillier,L., Lennon,G., Becker,M., Bonaldo,M.F., Chiapelli,B., Chissoe,S., Dietrich,N., DuBuque,T., Favello,A., Gish,W., Hawkins,M., Hultman,M., Kucaba,T., Lacy,M., Le,M., Le,N., Mardis,E., Moore,B., Morris,M., Parsons,J., Prange,C., Rifkin,L., Rohlfing,T., Schellenberg,K., Soares,M.B., Tan,F., Thierry-Meg,J., Trevaskis,E., Underwood,K., Wohldmann,P., Waterston,R., Wilson,R. and Marra,M.
TITLE Generation and analysis of 280,000 human expressed sequence tags
JOURNAL Genome Res. 6 (9), 807-828 (1996)
MEDLINE 97044478
COMMENT Contact: Wilson RK
Washington University School of Medicine
4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
Tel: 314 286 1800
Fax: 314 286 1810
Email: est@watson.wustl.edu

```
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            Trace considered overall poor quality
            Insert Length: 1438    Std Error: 0.00
            Seq primer: reverse ET
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .46
                      /organism="Homo sapiens"
                      /db_xref="GDB:3797462"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:248216"
                      /clone_lib="Soares fetal liver spleen 1NFLS"
                      /sex="male"
                      /dev_stage="20 week-post conception fetus"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: Liver and Spleen; Vector: pT7T3D (Pharmacia)
                      with a modified polylinker; Site_1: Pac I; Site_2: Eco RI;
                      1st strand cDNA was primed with a Pac I - oligo(dT) primer
                      [5' AACTGGAAGAATTAATTAAAGATCTTTTTTTTTTTTTTTTTT 3'],
                      double-stranded cDNA was ligated to Eco RI adaptors
                      (Pharmacia), digested with Pac I and cloned into the Pac I
                      and Eco RI sites of the modified pT7T3 vector.  Library
                      went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        6 a      20 c       7 g      12 t       1 others
ORIGIN

  Query Match              2.6%;  Score 45;  DB 14;  Length 46;
  Best Local Similarity   97.8%;  Pred. No. 0.78;
  Matches   45;  Conservative    0;  Mismatches    1;  Indels    0;  Gaps    0;

Qy   1700 ACTCTCTGCCTACCTGCCTGAGCCATGTTCACCTGCCCACTTGTCC 1745
           ||||||||||||| |||||||||||||||||||||||||||||||
Db      1 ACTCTCTGCCTACCNGCCTGAGCCATGTTCACCTGCCCACTTGTCC 46


RESULT 2
R38968/c
LOCUS       R38968                    28 bp    mRNA    linear   EST 04-MAY-1995
DEFINITION  yd07c08.s1 Soares infant brain 1NIB Homo sapiens cDNA clone
            IMAGE:25073 3' similar to gb:X66363 SERINE/THREONINE-PROTEIN KINASE
            PCTAIRE-1 (HUMAN);, mRNA sequence.
ACCESSION   R38968
VERSION     R38968.1  GI:796424
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 28)
  AUTHORS   Hillier,L., Clark,N., Dubuque,T., Elliston,K., Hawkins,M., Holman
            ,M., Hultman,M., Kucaba,T., Le,M., Lennon,G., Marra,M., Parsons,J.,
            Rifkin,L., Rohlfing,T., Soares,M., Tan,F., Trevaskis,E., Waterston
            ,R., Williamson,A., Wohldmann,P. and Wilson,R.
  TITLE     The WashU-Merck EST Project
  JOURNAL   Unpublished (1995)
COMMENT     Contact: Wilson RK
            Washington University School of Medicine
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: est@watson.wustl.edu
            Insert Size: 1349
            High quality sequence starts: 1 High quality sequence stops: 1
            Source: IMAGE Consortium, LLNL This clone is available royalty-free
            through LLNL ; contact the IMAGE Consortium (info@image.llnl.gov)
            for further information. Trace considered overall poor quality
            Insert Length: 1349    Std Error: 0.00
            Seq primer: -21m13
            High quality sequence stop: 1.
FEATURES              Location/Qualifiers
     source           1. .28
                      /organism="Homo sapiens"
                      /db_xref="GDB:397420"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:25073"
                      /clone_lib="Soares infant brain 1NIB"
                      /sex="female"
                      /dev_stage="73 days post natal"
                      /lab_host="DH10B (ampicillin resistant)"
                      /note="Organ: whole brain; Vector: Lafmid BA; Site_1: Not
                      I; Site_2: Hind III; 1st strand cDNA was primed with a Not
                      I - oligo(dT) primer [5'
                      AACTGGAAGAATTCGCGGCCGCAGGAATTTTTTTTTTTTTTTTT 3'];
                      double-stranded cDNA was ligated to Hind III adaptors
                      (Pharmacia), digested with Not I and directionally cloned
                      into the Not I and Hind III sites of the Lafmid BA vector.
                      Library went through one round of normalization. Library
                      constructed by Bento Soares and M.Fatima Bonaldo."
BASE COUNT        8 a      5 c      11 g      4 t
ORIGIN

  Query Match              1.6%;  Score 28;  DB 14;  Length 28;
  Best Local Similarity   100.0%;  Pred. No. 9.1e+03;
  Matches   28;  Conservative    0;  Mismatches    0;  Indels    0;  Gaps    0;

Qy   1716 CCTGAGCCATGTTCACCTGCCCACTTGT 1743
           ||||||||||||||||||||||||||||
Db     28 CCTGAGCCATGTTCACCTGCCCACTTGT 1


RESULT 3
AW247978
LOCUS       AW247978                  48 bp    mRNA    linear   EST 07-JAN-2000
DEFINITION  2819657.5prime NIH_MGC_7 Homo sapiens cDNA clone IMAGE:2819657 5',
            mRNA sequence.
ACCESSION   AW247978
VERSION     AW247978.1  GI:6591066
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   NIH-MGC http://mgc.nci.nih.gov/.
  TITLE     National Institutes of Health, Mammalian Gene Collection (MGC)
  JOURNAL   Unpublished (1999)
COMMENT     Other_ESTs: 2819657.3prime
            Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: DCTD/DTP cDNA Library Preparation: Ling
            Hong/Rubin Laboratory cDNA Library Arrayed by: The I.M.A.G.E.
            Consortium (LLNL) DNA Sequencing by: Berkeley MGC sequencing
            project Clone distribution: MGC clone distribution information can
            be found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html Base Calling / Quality
            Scores: PHRED from University of Washington Genome Center. Vector
            Trimming: cross_match from University of Washington Genome Center
            PHRAP suite. Poly-T Identification: patMatch.pl from Berkeley
            Drosophila Genome Project. University of Washington Genome Center:
            http://www.genome.washington.edu Low Quality Sequence: 8 contiguous
            PHRED high quality bases following vector sequence. Very Low
            Quality Sequence: Trace file contained 48 contiguous distinct peaks
            following vector sequence.
            Plate: LLCM2  row: C  column: 18
            High quality sequence stop: 8.
FEATURES              Location/Qualifiers
     source           1. .48
                      /organism="Homo sapiens"
                      /db_xref="taxon:9606"
                      /clone="IMAGE:2819657"
                      /clone_lib="NIH_MGC_7"
                      /tissue_type="small cell carcinoma"
                      /cell_line="MGC3"
                      /lab_host="DH10B (phage-resistant)"
```

```
                     /note="Organ: lung; Vector: pOTB7; Site_1: XhoI; Site_2:
                     EcoRI; cDNA made by oligo-dT priming.  Directionally
                     cloned into EcoRI/XhoI sites using the following 5'
                     adaptor: GGCACGAG(G).  Size-selected >500bp for average
                     insert size 1.8kb.  Library constructed by Ling Hong in
                     the laboratory of Gerald M. Rubin (University of
                     California, Berkeley) using  ZAP-cDNA synthesis kit
                     (Stratagene) and Superscript II RT (Life Technologies)."
BASE COUNT       11 a      14 c      17 g       6 t
ORIGIN

  Query Match             1.3%;  Score 23.2;  DB 10;  Length 48;
  Best Local Similarity   77.8%;  Pred. No. 1.8e+05;
  Matches   28;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy   1573 TCAGGCAGGCCAGCTTTCCGCGTGGTGGACACCGAG 1608
           ||||||| || ||| ||||| ||| | ||| ||||
Db     11 TCAGGCAATCCTGCTCTCCGCCTGGGGAACAGCGAG 46


RESULT 4
AU107934
LOCUS       AU107934                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU107934 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HSI01159, mRNA sequence.
ACCESSION   AU107934
VERSION     AU107934.1  GI:13557456
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HSI01159"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        7 a       9 c      25 g       9 t
ORIGIN

  Query Match             1.2%;  Score 21.4;  DB 9;  Length 50;
  Best Local Similarity   80.6%;  Pred. No. 5.1e+05;
  Matches   25;  Conservative    0;  Mismatches    6;  Indels    0;  Gaps    0;

Qy    227 AGAGTGGTGGTGGTGGCGGCAGTGACCCTGG 257
          ||||||| |||||||||||| | | |  |||
Db      1 AGAGTGGAGGTGGTGGCGGCGGCGGCTTTGG 31


RESULT 5
AZ346286
LOCUS       AZ346286                 36 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0081C01R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0081C01 R, DNA sequence.
ACCESSION   AZ346286
VERSION     AZ346286.1  GI:10425619
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 36)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0081  row: C  column: 01
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 36.
FEATURES             Location/Qualifiers
     source          1. .36
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0081C01"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        3 a       4 c      16 g      13 t
ORIGIN

  Query Match             1.2%;  Score 21.2;  DB 17;  Length 36;
  Best Local Similarity   76.5%;  Pred. No. 4.8e+05;
  Matches   26;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    212 AGATAGGCCTGGATGAGAGTGGTGGTGGTGGCGG 245
          ||||    |||  |||| ||||||||||||| ||
Db      1 AGATTTCTCTGCCTGAGTGTGGTGGTGGTGGTGG 34


RESULT 6
AU102877
LOCUS       AU102877                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU102877 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            COL03075, mRNA sequence.
```

```
ACCESSION   AU102877
VERSION     AU102877.1  GI:13552398
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="COL03075"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       11 a     20 c     10 g      9 t
ORIGIN

  Query Match             1.2%;  Score 21;  DB 9;  Length 50;
  Best Local Similarity   82.8%;  Pred. No. 6.3e+05;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    491 ACATCCGGCTGCCTGAGGGCTACCTGGAG 519
          |||||| ||||||||||  || |||| ||
Db      1 ACATCCAGCTGCCTGAGACCTTCCTGCAG 29


RESULT 7
AU105237/c
LOCUS       AU105237                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU105237 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            HRC08919, mRNA sequence.
ACCESSION   AU105237
VERSION     AU105237.1  GI:13554758
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="HRC08919"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT       13 a     10 c     22 g      5 t
ORIGIN

  Query Match             1.2%;  Score 21;  DB 9;  Length 50;
  Best Local Similarity   82.8%;  Pred. No. 6.3e+05;
  Matches   24;  Conservative    0;  Mismatches    5;  Indels    0;  Gaps    0;

Qy    558 CAGCCGCCGCCTCCGTCGTGTCAGCCTAT 586
          |||| |||||| || || |||| ||||||
Db     35 CAGCTGCCGCCGCCATCCTGTCCGCCTAT 7


RESULT 8
BM397711/c
LOCUS       BM397711                 50 bp    mRNA    linear   EST 17-JAN-2002
DEFINITION  5009-0-35-H11.t.2 Chilcoat/Turkewitz cDNA (large fraction)
            Tetrahymena thermophila cDNA, mRNA sequence.
ACCESSION   BM397711
VERSION     BM397711.1  GI:18197764
KEYWORDS    EST.
SOURCE      Tetrahymena thermophila.
  ORGANISM  Tetrahymena thermophila
            Eukaryota; Alveolata; Ciliophora; Oligohymenophorea;
            Hymenostomatida; Tetrahymenina; Tetrahymena.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Turkewitz,A.P., Karrer,K.M., Jahn,C., Orias,E., Kirk,K.E., Frankel
            ,J. and Klobutcher,L.
  TITLE     EST from Tetrahymena thermophila, strain CU428.1, growing cells
  JOURNAL   Unpublished (2002)
COMMENT     Contact: Turkewitz AP
            Molecular Genetics and Cell Biology
            University of Chicago
            920 E. 58th Street, Chicago, IL 60637, USA
            Tel: 773 702 4374
            Fax: 773 702 3172
            Email: apturkew@midway.uchicago.edu
            Seq primer: T3.
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Tetrahymena thermophila"
                     /strain="CU428.1"
                     /db_xref="taxon:5911"
                     /clone_lib="Chilcoat/Turkewitz cDNA (large fraction)"
                     /note="Vector: BlueScript2 SK+; Details on library
                     preparation can be found in Chilcoat and Turkewitz (2001)
                     Proc. Natl. Acad. Sci USA, 98: 8709-8713."
BASE COUNT       10 a      7 c     23 g      2 t      8 others
ORIGIN

  Query Match             1.2%;  Score 21;  DB 13;  Length 50;
  Best Local Similarity   60.0%;  Pred. No. 6.3e+05;
  Matches   27;  Conservative    0;  Mismatches   18;  Indels    0;  Gaps    0;

Qy     82 CCCCGCGGCTCTGAGGTTGCTCGCGCGCCCCCGCCGATCGCCATG 126
          ||||||| |||  || ||  || | | |       || || || ||
Db     50 CCCCGCGACTCCCAGCTTTGTCCCNCNCNNNNNNCGCTCTCCCTG 6


RESULT 9
AZ993993
LOCUS       AZ993993                 46 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0279E13F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0279E13 F, DNA sequence.
ACCESSION   AZ993993
VERSION     AZ993993.1  GI:13865220
KEYWORDS    GSS.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 46)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0279  row: E  column: 13
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 46.
FEATURES             Location/Qualifiers
     source          1. .46
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0279E13"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        8 a      0 c     23 g     15 t
ORIGIN

  Query Match             1.2%;  Score 20.8;  DB 17;  Length 46;
  Best Local Similarity   70.0%;  Pred. No. 6.8e+05;
  Matches   28;  Conservative    0;  Mismatches   12;  Indels    0;  Gaps    0;

Qy    223 GATGAGAGTGGTGGTGGTGGCGGCAGTGACCCTGGAGAGG 262
          |||||   |||||||||||| ||   |||   ||| || |
Db      3 GATGATGATGGTGGTGGTGGTGGTGATGATGATGGTGATG 42


RESULT 10
AZ311362
LOCUS       AZ311362                 48 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0026F16R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0026F16 R, DNA sequence.
ACCESSION   AZ311362
VERSION     AZ311362.1  GI:10354248
KEYWORDS    GSS.
SOURCE      house mouse.
```

```
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 48)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0026  row: F  column: 16
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 48.
FEATURES             Location/Qualifiers
     source          1. .48
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0026F16"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       18 a      9 c      6 g     15 t
ORIGIN

  Query Match             1.2%;  Score 20.6;  DB 17;  Length 48;
  Best Local Similarity   67.4%;  Pred. No. 7.8e+05;
  Matches   29;  Conservative    0;  Mismatches   14;  Indels    0;  Gaps    0;

Qy   1499 CTACTTCCATATTTGCACTAAAGGAGATTCAGCTACAAAAGGA 1541
          |||  | ||| |||  ||| ||| ||| ||| | |  ||| ||
Db      5 CTAAATGCATGTTTCTACTGAAGTAGAATCACCAATTAAAAGA 47


RESULT 11
AU106960/c
LOCUS       AU106960                 50 bp    mRNA    linear   EST 30-AUG-2001
DEFINITION  AU106960 Sugano Homo sapiens cDNA library Homo sapiens cDNA clone
            CAS09689, mRNA sequence.
ACCESSION   AU106960
VERSION     AU106960.1  GI:13556481
KEYWORDS    EST.
SOURCE      human.
```

```
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 50)
  AUTHORS   Suzuki,Y., Taira,H., Tsunoda,T., Mizushima-Sugano,J., Sese,J., Hata
            ,H., Ota,T., Isogai,T., Tanaka,T., Morishita,S., Okubo,K., Sakaki
            ,Y., Nakamura,Y., Suyama,A. and Sugano,S.
  TITLE     Diverse transcriptional initiation revealed by fine, large-scale
            mapping of mRNA start sites
  JOURNAL   EMBO Rep. 2 (5), 388-393 (2001)
  MEDLINE   21270072
COMMENT     Contact: Yutaka Suzuki
            Department of Virology
            Institute of Medical Science, University of Tokyo
            4-6-1, Shirokanedai, Minatoku, Tokyo 108-8639, Japan
            Email: ysuzuki@ims.u-tokyo.ac.jp
            Suzuki,Y., Yoshitomo-Nakagawa,K., Maruyama,K., Suyama,A. and Sugano
            ,S. Construction and characterization of a full length-enriched and
            a 5'-end-enriched cDNA library. Gene 200 (1-2), 149-156 (1997).
FEATURES             Location/Qualifiers
     source          1. .50
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="CAS09689"
                     /clone_lib="Sugano Homo sapiens cDNA library"
                     /note="Differential display comparison of untreated and
                     dimethylfumarate treated U937 cells"
BASE COUNT        7 a     14 c     18 g     11 t
ORIGIN

  Query Match              1.2%;  Score 20.6;  DB 9;  Length 50;
  Best Local Similarity   74.3%;  Pred. No. 7.9e+05;
  Matches   26;  Conservative    0;  Mismatches    9;  Indels    0;  Gaps    0;

Qy    922 CTGTTCCAGCTGCTCCGTGGCCTGGCCTACTGCCA 956
          || | |||||||| || |||||  |||| | || |
Db     43 CTTTCCCAGCTGCGCCATGGCCGCGCCTGCCGCAA 9


RESULT 12
AZ331536/c
LOCUS       AZ331536                 47 bp    DNA     linear   GSS 29-SEP-2000
DEFINITION  1M0059H04R Mouse 10kb plasmid UUGC1M library Mus musculus genomic
            clone UUGC1M0059H04 R, DNA sequence.
ACCESSION   AZ331536
VERSION     AZ331536.1  GI:10394326
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 47)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0059  row: H  column: 04
            Seq primer: CACACAGGAAACAGCTATGACC
            Class: plasmid ends
            High quality sequence stop: 47.
FEATURES             Location/Qualifiers
     source          1. .47
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC1M0059H04"
                     /clone_lib="Mouse 10kb plasmid UUGC1M library"
                     /sex="Male"
                     /lab_host="E. Coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (male) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT        5 a     12 c     20 g     10 t
ORIGIN

  Query Match              1.2%;  Score 20.4;  DB 17;  Length 47;
  Best Local Similarity   71.1%;  Pred. No. 8.6e+05;
  Matches   27;  Conservative    0;  Mismatches   11;  Indels    0;  Gaps    0;

Qy   1644 GCTGGAGGGATGCCACACCCCTCACAGGGCAGCCCCCA 1681
          | | || ||  ||| | ||| ||||| |||  ||||||
Db     38 GTTTGAAGGCGGCCCCCCCCATCACAAGGCCCCCCCCA 1


RESULT 13
AZ985975
LOCUS       AZ985975                 45 bp    DNA     linear   GSS 27-APR-2001
DEFINITION  2M0268F01F Mouse 10kb plasmid UUGC2M library Mus musculus genomic
            clone UUGC2M0268F01 F, DNA sequence.
ACCESSION   AZ985975
VERSION     AZ985975.1  GI:13857202
KEYWORDS    GSS.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 45)
  AUTHORS   Dunn,D., Aoyagi,A., Barber,M., Beacorn,T., Duval,B., Hamil,C.,
            Islam,H., Longacre,S., Mahmoud,M., Meenen,E., Pedersen,T., Reilly
            ,M., Rose,M., Rose,R., Stokes,R., Tingey,A., von Niederhausern,A.
            and Wright,D.,Weiss,R.
  TITLE     Mouse whole genome scaffolding with paired end reads from 10kb
            plasmid inserts
  JOURNAL   Unpublished (2000)
COMMENT     Contact: Robert B. Weiss
            University of Utah Genome Center
            University of Utah
            Rm. 308, Biomedical Polymers Research Bldg., 20 S. 2030 E., SLC, UT
            84112, USA
            Tel: 801 585 5606
            Fax: 801 585 7177
            Email: ddunn@genetics.utah.edu
            Insert Length: 10000   Std Error: 0.00
            Plate: 0268  row: F  column: 01
            Seq primer: CGTTGTAAAACGACGGCCAGT
            Class: plasmid ends
            High quality sequence stop: 45.
FEATURES             Location/Qualifiers
```

```
     source          1. .45
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="UUGC2M0268F01"
                     /clone_lib="Mouse 10kb plasmid UUGC2M library"
                     /sex="Female"
                     /lab_host="E. coli strain XL10-Gold, T1-resistant, F-"
                     /note="Vector: PWD42nv; Purified genomic DNA from M.
                     musculus C57BL/6J (female) was obtained  from the Jackson
                     Laboratory Mouse DNA Resource
                     (http://www.jax.org/resources/documents/dnares/). The DNA
                     was hydrodynamically sheared by repeated passage through a
                     0.005 inch orifice at constant velocity. The sheared DNA
                     was blunt end-repaired with T4 DNA polymerase and T4
                     polynucleotide kinase. Adaptor oligonucleotides were
                     ligated to the blunt ends in high  molar excess. The
                     adaptored DNA was purified and size-selected for a 9.5 to
                     10.5 kb range using preparative agarose gel
                     electrophoresis. Vector DNA was prepared from a derivative
                     of pWD42 (gi|4732114|gb|AF129072.1), a copy-number
                     inducible derivative of plasmid R1. The vector was ligated
                     with adaptors complementary to the insert adaptors and
                     purified. The sheared, adaptored mouse DNA was annealed to
                     adaptored vector DNA, and transformed into
                     chemically-competent E. coli XL10-Gold (Stratagene) cells
                     and selected for ampicillin resistance."
BASE COUNT       18 a      0 c      24 g      3 t
ORIGIN

  Query Match             1.2%;  Score 20.2;  DB 17;  Length 45;
  Best Local Similarity   68.3%;  Pred. No. 9.5e+05;
  Matches   28;  Conservative    0;  Mismatches   13;  Indels    0;  Gaps    0;

Qy       4 AAGCAGCGTAAAGGATGGACAGGAATGCAGAGGTAGGCAGG 44
           ||  |  | |||||  ||| |||||||  |||| |||  ||
Db       4 AAAGAAAGAAAAGGGGGGAGAGGAATGGGGAGGGAGGGTGG 44


RESULT 14
AA204601/c
LOCUS       AA204601                 49 bp    mRNA    linear   EST 19-FEB-1997
DEFINITION  mu25c05.r1 Soares_thymus_2NbMT Mus musculus cDNA clone IMAGE:640424
            5' similar to SW:143G_BOVIN P29359 14-3-3 PROTEIN GAMMA ;, mRNA
            sequence.
ACCESSION   AA204601
VERSION     AA204601.1  GI:1800608
KEYWORDS    EST.
SOURCE      house mouse.
  ORGANISM  Mus musculus
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Rodentia; Sciurognathi; Muridae; Murinae; Mus.
REFERENCE   1  (bases 1 to 49)
  AUTHORS   Marra,M., Hillier,L., Allen,M., Bowles,M., Dietrich,N., Dubuque,T.,
            Geisel,S., Kucaba,T., Lacy,M., Le,M., Martin,J., Morris,M.,
            Schellenberg,K., Steptoe,M., Tan,F., Underwood,K., Moore,B.,
            Theising,B., Wylie,T., Lennon,G., Soares,B., Wilson,R. and
            Waterston,R.
  TITLE     The WashU-HHMI Mouse EST Project
  JOURNAL   Unpublished (1996)
COMMENT     Contact: Marra M/Mouse EST Project
            WashU-HHMI Mouse EST Project
            Washington University School of MedicineP
            4444 Forest Park Parkway, Box 8501, St. Louis, MO 63108
            Tel: 314 286 1800
            Fax: 314 286 1810
            Email: mouseest@watson.wustl.edu
            This clone is available royalty-free through LLNL ; contact the
            IMAGE Consortium (info@image.llnl.gov) for further information.
            MGI:392416
            Trace considered overall poor quality
            Possible reversed clone: similarity on wrong strand
            Seq primer: -28M13 rev2 from Amersham
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .49
                     /organism="Mus musculus"
                     /strain="C57BL/6J"
                     /db_xref="taxon:10090"
                     /clone="IMAGE:640424"
                     /clone_lib="Soares_thymus_2NbMT"
                     /sex="male"
                     /tissue_type="Thymus"
                     /dev_stage="4 weeks"
                     /lab_host="DH10B"
                     /note="Vector: pT7T3D-Pac (Pharmacia) with a modified
                     polylinker; Site_1: Not I; Site_2: Eco RI; 1st strand cDNA
                     was primed with a Not I - oligo(dT) primer [5'
                     TGTTACCAATCTGAAGTGGGAGCGGCCGCGTTTTTTTTTTTTTTTTTTTTTTTT
                     3']; double-stranded cDNA was ligated to Eco RI adaptors
                     (Pharmacia), digested with Not I and cloned into the Not I
                     and Eco RI sites of the modified pT7T3 vector. RNA
                     provided by Dr. Bertrand Jordan. Library went through two
                     rounds of normalization, and was constructed by Bento
                     Soares and M.Fatima Bonaldo."
BASE COUNT       14 a      23 c       5 g       7 t
ORIGIN

  Query Match             1.2%;  Score 20.2;  DB 9;  Length 49;
  Best Local Similarity   75.8%;  Pred. No. 9.9e+05;
  Matches   25;  Conservative    0;  Mismatches    8;  Indels    0;  Gaps    0;

Qy    1156 ATGTGGGGTGTGGGCTGCATCTTCTATGAGATG 1188
           ||| || ||||| ||||||| | || || | ||
Db      33 ATGGGGTGTGTGAGCTGCATGTGCTCTGTGCTG 1


RESULT 15
AI475974/c
LOCUS       AI475974                 40 bp    mRNA    linear   EST 14-APR-1999
DEFINITION  tl96b06.x1 NCI_CGAP_Co14 Homo sapiens cDNA clone IMAGE:2154899 3'
            similar to TR:Q17352 Q17352 RRM-TYPE RNA BINDING PROTEIN. ;contains
            element MSR1 repetitive element ;, mRNA sequence.
ACCESSION   AI475974
VERSION     AI475974.1  GI:4329019
KEYWORDS    EST.
SOURCE      human.
  ORGANISM  Homo sapiens
            Eukaryota; Metazoa; Chordata; Craniata; Vertebrata; Euteleostomi;
            Mammalia; Eutheria; Primates; Catarrhini; Hominidae; Homo.
REFERENCE   1  (bases 1 to 40)
  AUTHORS   NCI-CGAP http://www.ncbi.nlm.nih.gov/ncicgap.
  TITLE     National Cancer Institute, Cancer Genome Anatomy Project (CGAP),
            Tumor Gene Index
  JOURNAL   Unpublished (1997)
COMMENT     Contact: Robert Strausberg, Ph.D.
            Email: cgapbs-r@mail.nih.gov
            Tissue Procurement: Christopher Moskaluk, M.D., Ph.D., Michael R.
            Emmert-Buck, M.D., Ph.D.
            cDNA Library Preparation: Life Technologies, Inc.
            cDNA Library Arrayed by: Greg Lennon, Ph.D.
            DNA Sequencing by: Washington University Genome Sequencing Center
            Clone distribution: NCI-CGAP clone distribution information can be
            found through the I.M.A.G.E. Consortium/LLNL at:
            www-bio.llnl.gov/bbrp/image/image.html

            Trace considered overall poor quality
            Insert Length: 416   Std Error: 0.00
            Seq primer: -40UP from Gibco
            High quality sequence stop: 1.
FEATURES             Location/Qualifiers
     source          1. .40
                     /organism="Homo sapiens"
                     /db_xref="taxon:9606"
                     /clone="IMAGE:2154899"
                     /clone_lib="NCI_CGAP_Co14"
```

/tissue_type="moderately-differentiated adenocarcinoma"
/lab_host="DH10B"
/note="Organ: colon; Vector: pCMV-SPORT6; Site_1: SalI;
Site_2: NotI; Cloned unidirectionally. Primer: Oligo dT.
Average insert size 1.7 kb. Life Technologies catalog #:
11531-019"
BASE COUNT 9 a 13 c 14 g 4 t
ORIGIN

Query Match 1.1%; Score 20; DB 9; Length 40;
Best Local Similarity 72.2%; Pred. No. 1e+06;
Matches 26; Conservative 0; Mismatches 10; Indels 0; Gaps 0;

```
Qy     232 GGTGGTGGTGGCGGCAGTGACCCTGGAGAGGCCCCC 267
           |||||||||| |   |   ||||||| ||  |||||
Db      39 GGTGGTGGTGTCTTTACCAACCCTGGTGACCCCCCC 4
```

Search completed: March 4, 2003, 00:50:32
Job time : 2635 secs